DOCUMENT RESUME

ED 044 273 SE 009 210

AUTHOR	Trigg, George L.
TITLE	Crucial Experiments in Quantum Physics.
INSTITUTION	Commission on Coll. Physics, College Park, Md.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	66
NOTE	42p.; Monograph written for the Conference on the New Instructional Materials in Physics (University of Washington, Seattle, 1965)
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.20
DESCRIPTORS	*College Science, *Experiments, Instructional Materials, *Physics, *Quantum Mechanics, Resource Materials, *Science History

ABSTRACT

The six experiments included in this monography are titled Blackbody Radiation, Collision of Electrons with Atoms, The Photoelectric Effect, Magnetic Properties of Atoms, The Scattering of X-Rays, and Diffraction of Electrons by a Crystal Lattice. The discussion provides historical background by giving description of the original experiments and events which contributed to them. Many original figures are given and quotes from original accounts are presented. An attempt was made to present the essentials of the method of each experiment and the difficulties of execution and interpretation encountered. (PR)

EDO 44273

FILE COPY

Crucial Experiments in Quantum Physics

GEORGE L. TRIGG

Brookhaven National Laboratory

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SE 009 210

© 1966 The University of Washington, Seattle

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Ronald Geballe
University of Washington
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

ED 044273

GENERAL PREFACE

This monograph was written for the Conference on the New Instructional Materials in Physics, held at the University of Washington in the summer of 1965. The general purpose of the conference was to create effective ways of presenting physics to college students who are not preparing to become professional physicists. Such an audience might include prospective secondary school physics teachers, prospective practitioners of other sciences, and those who wish to learn physics as one component of a liberal education.

At the Conference some 40 physicists and 12 filmmakers and designers worked for periods ranging from four to nine weeks. The central task, certainly the one in which most physicists participated, was the writing of monographs.

Although there was no consensus on a single approach, many writers felt that their presentations ought to put more than the customary emphasis on physical insight and synthesis. Moreover, the treatment was to be "multi-level" --- that is, each monograph would consist of several sections arranged in increasing order of sophistication. Such papers, it was hoped, could be readily introduced into existing courses or provide the basis for new kinds of courses.

Monographs were written in four content areas: Forces and Fields, Quantum Mechanics, Thermal and Statistical Physics, and the Structure and Properties of Matter. Topic selections and general outlines were only loosely coordinated within each area in order to leave authors free to invent new approaches. In point of fact, however, a number of monographs do relate to others in complementary ways, a result of their authors' close, informal interaction.

Because of stringent time limitations, few of the monographs have been completed, and none has been extensively rewritten. Indeed, most writers feel that they are barely more than clean first drafts. Yet, because of the highly experimental nature of the undertaking, it is essential that these manuscripts be made available for careful review

by other physicists and for trial use with students. Much effort, therefore, has gone into publishing them in a readable format intended to facilitate serious consideration.

So many people have contributed to the project that complete acknowledgement is not possible. The National Science Foundation supported the Conference. The staff of the Commission on College Physics, led by E. Leonard Jossem, and that of the University of Washington physics department, led by Ronald Geballe and Ernest M. Henley, carried the heavy burden of organization. Walter C. Michels, Lyman G. Parratt, and George M. Volkoff read and criticized manuscripts at a critical stage in the writing. Judith Bregman, Edward Gerjuoy, Ernest M. Henley, and Lawrence Wilets read manuscripts editorially. Martha Ellis and Margery Lang did the technical editing; Ann Widditsch supervised the initial typing and assembled the final drafts. James Grunbaum designed the format and, assisted in Seattle by Roselyn Pape, directed the art preparation. Richard A. Mould has helped in all phases of readying manuscripts for the printer. Finally, and crucially, Jay F. Wilson, of the D. Van Nostrand Company, served as Managing Editor. For the hard work and steadfast support of all these persons and many others, I am deeply grateful.

Edward D. Lambe
Chairman, Panel on the
New Instructional Materials
Commission on College Physics

C R U C I A L E X P E R I M E N T S I N Q U A N T U M P H Y S I C S

PREFACE

This work serves a double function. On the one hand, it provides some of the historical background of quantum theory for the student who has already achieved some understanding of the theory; on the other hand, it helps to make easier for the beginner the acceptance of some of the strange-seeming concepts of quantum physics. In general, I have attempted to present each experiment from the point of view of the average physicist of the time; I have made specific mention of the one conscious exception of this rule.

The list of experiments that are discussed is naturally somewhat arbitrary. I have deliberately omitted some that pertained to atomism but not to quantum theory, such as Millikan's work on the electronic charge. I have also ignored some which, while their conclusions were essential to later developments in quantum theory, nevertheless were themselves primarily directed at questions of atomic structure; the experiment of Geiger and Marsden is an outstanding example. However, I have tried to make the discussion of each experiment reasonably complete, in the sense that the essentials of the method and the difficulties both of execution and of interpretation are made clear.

I have benefited greatly from criticisms from Dr. Walter Michels, Dr. Lawrence Wilets, and Mr. Roger Atlas.

George L. Trigg

C O N T E N T S

	PREFACE	
1	INTRODUCTION	1
2	BLACKBODY RADIATION	2
3	COLLISIONS OF ELECTRONS WITH ATOMS	10
4	THE PHOTOELECTRIC EFFECT	14
5	MAGNETIC PROPERTIES OF ATOMS	21
6	THE SCATTERING OF X RAYS	25
7	DIFFRACTION OF ELECTRONS BY A CRYSTAL LATTICE	30

1 INTRODUCTION

It is natural human behavior to expect that regularities which hold under a fairly wide range of conditions will continue to hold under all conditions. Thus, for example, it would be futile to send a rocket toward Mars carrying equipment with which to take pictures of the surface of that planet unless we were at least reasonably confident that the equipment, and the rocket itself, would behave near Mars in accordance with the same laws we have discovered on the surface of the earth. Similarly, we have faith that an automobile designed five years ago in the light of laws as they were known then will still operate properly tomorrow.

But confidence in such extensions, or extrapolations, is sometimes misplaced. For example, a famous law of physics, discovered in 1662 by Robert Boyle, states that the product of the volume of a particular quantity of gas and the pressure it exerts on the walls of its container is constant as long as the temperature of the gas remains unchanged, any increase in one of the two quantities being exactly compensated by a decrease in the other. If the temperature is too low, however, or the pressure too high, the law fails; the gas starts to condense, and there may be a substantial change of volume with no compensating change of pressure. The failure of an extrapolation is commonly linked with the existence of laws or phenomena not envisioned in the original relationship. Thus the failure of Boyle's law is related to the fact that gases can condense to form liquids.

The branch of physics known as quantum theory has at its basis the results of a rather small number of experiments, most of which revealed such failures of extrapolation. (In at least one case the failure was of a different sort: The previous theory simply gave no information.) They have one feature in common, namely, that the additional concepts that must be considered in order to provide a coherent explanation are foreign to our ordinary experience. These novel concepts are fundamental to quantum theory, and it is their strangeness that usually causes the greatest difficulty in understanding the theory.

The present work consists of a description and brief discussion of several of these experiments as they were originally carried out, including many of the original figures and often quoting from the original accounts.

It will be noted that all these experiments deal either with sizes far removed from ordinary experience or with subtle details of the radiation of light and heat. The key to an appreciation of them may be described, therefore, as a refusal to be restrained by "common sense" outside the realm where common sense has been acquired.

2 BLACKBODY RADIATION

Historically, quantum physics originated in an attempt to give a complete description of the radiation from a black body. In order to see how this came about, it is first necessary to understand what a black body is and what its properties are.

Whenever light strikes the surface of an object, two effects occur: Some of the light is reflected from the surface, and some passes through the surface into the body of the object. The latter protion, in turn, undergoes at least one and possibly two further processes: Some (or all) of it is absorbed; some may reach another surface of the object and pass out, or be transmitted. We see an object, unless it is intrinsically luminous, only because it reflects some light into our eyes. Even the sky is visible because of light scattered toward our eyes. The less light an object reflects, the darker it appears to be. If an object should absorb all the light that fell on it, it would reflect none, and would appear perfectly black.[1] An object of this sort is called an ideal black body. No such body actually exists. Nevertheless, it is quite possible - and useful! - to act as though it did, and to determine many of the properties it would have; this sort of idealization is quite common in science.

Of course, the radiation that is absorbed by a black body carries energy. The internal energy of the body would be increased and the temperature of the body would rise indefinitely if there were no mechanism by which the body disposed of some energy. The mechanism is, simply, that the body radiates; in fact, not only a black body but any object left to itself in unchanging surroundings tends toward a state of equilibrium. When it has reached the equilibrium state, it radiates as much energy per unit time - as much power - as it absorbs.[2]

The reason for interest in the properties of a black body is somewhat abstruse. It begins with the fact that any object gives off radiation, at a rate which increases strongly with the temperature of the body. The wavelengths of the radiation range continuously over the spectrum, not only the visible portion but also the ultraviolet, the infrared, and all other parts; and the way in which the energy is distributed in wavelength also changes with the temperature of the body, as well as being dependent on the nature of the body. The measurement of this distribution is called the spectral emittance, designated by E_λ and defined as follows: The spectral emittance at wavelength λ is the energy associated with wavelengths around λ radiated per unit time, per unit surface area, per unit wavelength. Thus, $E_\lambda\, dt\, dS\, d\lambda$ is the energy in the wavelength range λ to $\lambda + d\lambda$ radiated by a surface element of area dS in a time interval dt. Secondly, as has been mentioned, any real object absorbs a fraction of the radiation striking it. The fraction absorbed depends on the nature of the body and on the wavelength of the incident light; let us denote the fraction at wavelength λ by A_λ. Both A_λ and E_λ may vary with temperature. There is a relationship between them, however, deduced by G. R. Kirchhoff in introducing the concept of a black body in

[1] Actually, of course, we would not be able to see the object at all, but only a chunk of space from which no light reached us.

[2] It must be emphasized that this is not the same as reflection. The characteristics of reflected light are determined partly by the properties of the reflecting body but also partly by those of the incident light. Radiation, on the other hand, is not in any way affected by incident light.

1860: The ratio between E_λ and A_λ for an arbitrary body at any given wavelength and temperature is equal to the spectral emittance of a black body at that wavelength and temperature. In symbols, if e_λ is the spectral emittance of a black body, the relationship is

$$E_\lambda/A_\lambda = e_\lambda.$$

O. Lummer and E. Pringsheim, whose experimental studies of blackbody radiation are to be described a little later, go on as follows[3]:

> If one knows, therefore, the radiation of a black body as a function of wavelength and temperature, then one knows thereby the laws of radiation for all those bodies whose absorptive power is likewise given as a function of wavelength and temperature. Experimentally, the reverse process is likely to be simpler, by a study of the radiation of a body to explain the absorption A with the help of the knowledge of e.

Evidently, knowledge of e_λ was desirable, and by the end of the nineteenth century the problem had already received considerable attention. Most of the efforts were empirical deductions from observations on real bodies. In 1896, F. Paschen, reporting his own results of this type in the journal Annalen der Physik, cited some half-dozen earlier works; Paschen's expression, $e_\lambda = C\lambda^{-a}e^{-c/\lambda t}$, where c, C, and a are constants for a given material, was one of the simplest. The one empirical result that has retained its validity was one given in the Sitzungsberichte der königliche Gesellschaft der Wissenschaft zu Wien, in 1879, by J. Stefan, that the total emittance, the total power (regardless of wavelength) radiated per unit area, is proportional to the fourth power of the absolute temperature of the radiating body.

Theoretical studies were not lacking. They were immensely simplified by the following property, also proved by Kirchhoff: If a region of space is bounded by material of whatever nature, provided only that it is not a mirror-like reflector, and these boundary walls are maintained at a uniform temperature T, the space will become filled with radiation identical in every respect with that emitted by a black body whose temperature is T. The method of proof is to show that otherwise it would be possible to bring about a violation of the second law of thermodynamics. As we shall see, the possibility of experimental measurement of e_λ rests upon essentially this property of a cavity.

In 1884, Ludwig Boltzmann published in Annalen der Physik two papers which together provided a proof that Stefan's empirical relationship must hold for a black body. (As a consequence, the relationship has become known as the Stefan-Boltzmann law.) His arguments involved a cavity whose walls were not at a uniform temperature, and which was subdivided by movable pistons with shutters in them; he applied the second law of thermodynamics to the transfer of energy from one wall to another by means of various motions of the pistons and operation of the shutters, making use of the fact that the radiation exerts a pressure on the pistons so that work is expended or absorbed in their motion. Nine years later, Willy Wien obtained two equally important results, which he published in the Sitzungsberichte der königliche Preussische Akademie der Wissenschaften zu Berlin. He noted that if a cavity were reduced in volume, the energy per unit volume in the cavity would be increased not only by the confinement to a smaller volume but also by virtue of the work done against radiation pressure. The density can also be increased by an increase of temperature, and the second law of thermodynamics relates the

[3]Translated by G. L. T.

increases produced in the two ways. The relationship must hold not only over all, but also for the energy density associated with every infinitesimal range of wavelengths. But the moving walls produce a change in wavelength of the reflected radiation, by the Doppler effect; and so the temperature change must also alter the wavelength distribution. The quantitative consequences of this were, first, that the wavelength λ' at temperature T' that properly corresponds to the wavelength λ at temperature T is given by

$$\lambda'T' = \lambda T; \qquad (2.1)$$

second, that the spectral emittances for corresponding wavelengths at different temperatures are related by

$$e_\lambda/e_{\lambda'} = T^5/T'^5. \qquad (2.2)$$

In particular, if e_λ has a maximum value $e_{\lambda,max}$ at some wavelength λ_m, then $e_{\lambda,max}$ satisfies the relation

$$e_{\lambda,max}T^{-5} = \text{const.}, \qquad (2.2a)$$

while λ_m satisfies the relation

$$\lambda_m T = \text{const.} \qquad (2.1a)$$

Equations (2.1) and (2.2) are known as Wien's displacement laws. Together, although Wien did not note it, they imply that the expression for e_λ must be of the general form

$$e_\lambda = \lambda^{-5}f(\lambda T), \qquad (2.3)$$

or, equivalently,

$$e_\lambda = T^5F(\lambda T),$$

where $F(\lambda T) = (\lambda T)^{-5}f(\lambda T)$. The specific form of the function $f(\lambda T)$ or $F(\lambda T)$ is not determined by these arguments; merely the fact that they depend on λ and T only through the product λT. This was the most that could be established on the basis of classical theory without the addition of more or less questionable hypotheses.

Wien went on, in a later paper (1896) in Annalen der Physik, to obtain an explicit form for e_λ by making assumptions about the process of radiation from a molecule. His result was

$$e_\lambda = C\lambda^{-5}e^{-c/\lambda T}, \qquad (2.4)$$

in agreement with Eq. (2.3) and in support of Paschen's empirical results; but some physicists regarded his assumptions as dubious.

One other theoretical attempt deserves mention, although it actually postdated the experimental work. In 1900, Lord Rayleigh approached the problem in the following way: The radiation in a cavity at equilibrium must consist of standing waves. It is possible to calculate the number of different standing-wave modes of wavelengths between λ and $\lambda + d\lambda$ that can exist per unit volume of the cavity. According to classical statistical considerations, each of these modes should have the same average energy, namely an amount kT where T is the absolute temperature and k is a universal constant. These considerations lead to an expression for the energy per unit volume in the cavity per unit wavelength range, and thence to an equation for the spectral emittance,

$$e_\lambda = 8\pi ckT\lambda^{-4}.$$

This equation agrees with Eq. (2.3). However, it has one serious defect: It yields an infinite value for the total energy in the cavity.[4] The infinite value is due to larger and larger contributions from shorter and shorter wavelengths. Since short wavelengths are associated with ultraviolet radiation, the divergence is often referred to as the "ultraviolet catastrophe." (In the light of modern

[4] It is interesting to note that there were previous indications of the failure of the law of equipartition of energy. Maxwell had noticed as early as 1859 that the law was incapable of providing an adequate explanation of the ratio of specific heats of gases.

knowledge, a more appropriate name might be the "gamma-ray catastrophe.")

The experimental study was made possible by a slight modification of the discussion that gave the equivalence of radiation in a closed cavity and blackbody radiation. Suppose a cavity of the type envisioned earlier has a hole drilled through one wall. The hole must have a diameter large compared with its length, but its area must be small compared with that of the cavity walls. Under these circumstances, any radiation that falls on the hole is almost certain to pass into the cavity, there to be trapped by continual reflection back and forth by the walls, with some absorption at each reflection, so that the hole is in that respect a good approximation to a black body. Moreover, the radiation coming out of the hole will be a representative sample of the radiation in the cavity, which has already been described as equivalent to blackbody radiation.

The definitive experimental work was carried out by Lummer and Pringsheim in Charlottenburg, Germany. The first step, reported in Annalen der Physik in 1897, was to verify the Stefan-Boltzmann relationship. The procedure is simple in principle - merely to measure the energy radiated from a hole in a cavity maintained at a constant known temperature - but it required considerable care in its execution.

Two cavities were used, a copper one for temperatures up to 877°K and an iron one for temperatures from 799°K to 1561°K. The copper cavity was immersed in a molten mixture of sodium nitrate and potassium nitrate; the temperature of the bath could be held constant to within one or two degrees for as long as half an hour by control of the supply of gas to the heating flame. The iron cavity was heated by means of a special double-walled oven shown in Fig. 2.1. The hot gases from the flame passed around the cavity inside the inner wall of the oven, then between the two walls, and then

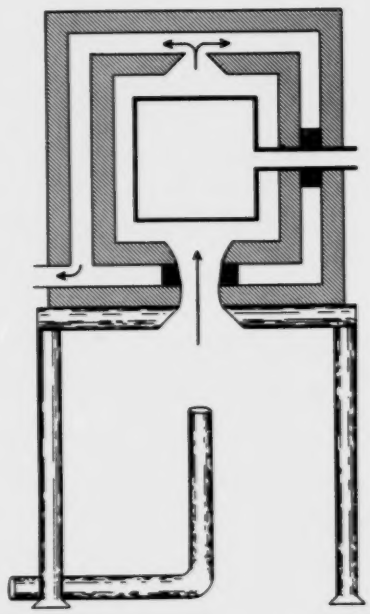

Fig. 2.1 Diagram of the double-walled oven as used for heating the iron cavity.

into the chimney flue. Temperatures up to 755°K were measured by means of mercury thermometers; higher temperatures, by a thermocouple.

The radiant power was measured by means of a bolometer. In this device, radiation falls on one of two blackened platinum wires and is absorbed, raising the temperature of the wire and therefore its electrical resistance. The increase in resistance is measured by comparison with the resistance of the other wire. Extensive precautions were taken to ensure that the energy recorded came only from the cavity, and to correct for possible variations in the fraction of radiation absorbed in the air along the path from cavity to sensing element. In fact, the only difficulty that was not almost entirely overcome was that of achieving truly uniform temperatures in the iron cavity. The final conclusion, based on observations over a range of temperature whose extremes differed by a factor of four, was that the Stefan-Boltzmann law is valid.

Lummer and Pringsheim then proceeded to a study of how the emissivity varies with wavelength at a given temperature. The results were reported in a series of three papers in the Verhandlungen der Deutschen physikalischen Gesellschaft in 1899 and 1900.

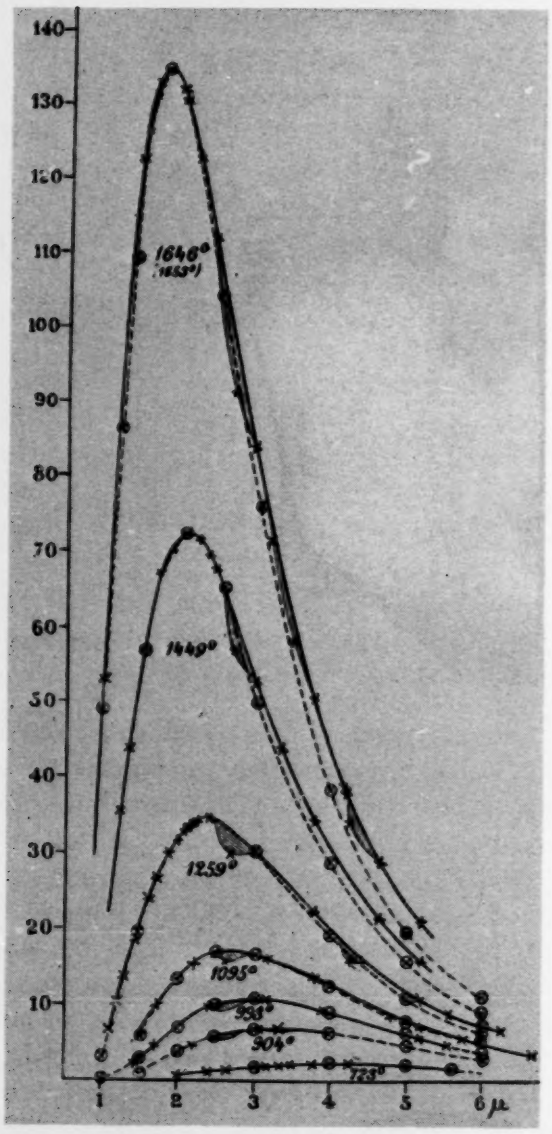

Fig. 2.2 Comparison of Lummer and Pringsheim's experimental data (plain crosses and solid curves), with Wien's formula (circled crosses and dashed curves). The shaded areas show the absorption by water vapor and carbon dioxide in the air.

Again the basic concept was simple and the basic procedure straightforward, and only the various precautions and corrections were complicated. Various cavities were used, at temperatures from 85 to about 1800°K. The lower temperatures were achieved by immersion in liquid air (85°K), boiling water (373°K), and molten saltpeter (around 600°K, depending on exact composition). Higher temperatures, up to about 1800°K, were obtained by electrical heating. At such temperatures, by far the largest fraction of the radiation is in the infrared region of the spectrum; the range of wavelengths studied was from about 1 micron to about 18 microns (visible light covers the range of wavelengths from about 0.4 to about 0.7 micron). A substantial difficulty in this spectral region is that water vapor and carbon dioxide, both normally present in the atmosphere, absorb strongly near certain wavelengths, especially around 1.8, 2.7, and 4.5 microns. In the earliest work, Lummer and Pringsheim merely attempted to correct for this. Later, they enclosed the spectrometer and bolometer in a container in which the air was dried and chemically purged of carbon dioxide, so that the necessary correction was greatly reduced. As in the work on the Stefan-Boltzmann law, strict precautions were taken to ensure that only the radiation of interest fell on the bolometer.

One method of presenting the results is simply as a curve showing e_λ as a function of λ for various temperatures as in Fig. 2.2. From such curves, Lummer and Pringsheim determined both the wavelength at which e_λ was maximum, and the maximum value, for testing Eqs. (2.1a) and (2.2a). It was simply a matter of seeing whether the appropriate combinations of factors were indeed constants. Already in their first report, they could make this statement[5]:

> It can therefore be regarded as proven by this series of observations that for the radiating body employed the maximum energy increases with the fifth power of the absolute temperature. Also the equation $\lambda_m T = A$ can be considered proven, since the deviations of the values of A from the average value lie within the observational errors possible from the determination of λ_m.

It is interesting to note that Lummer and Pringsheim repeated these tests on each series of observations

[5]Translated by G. L. T.

they carried out, for, as they put it, "The fulfilment of these three laws [the third being Stefan's law] is the conditio sine qua non if one wishes to draw from the radiation measurements any conclusion whatever about the form of the spectral equation (energy curve)." In fact, so firmly convinced were they of the truth of this statement that they discarded one series of observations because the maximum value of e_λ increased as $T^{5.2}$ rather than as T^5.[6]

Figure 2.2 shows, together with one set of experimental plots of the sort just discussed, the curves represented by Eq. (2.4) for the values of T used. The agreement between theory and experiment looks fair, but Lummer and Pringsheim were not satisfied, and devised a means of making a more sensitive test. If one takes the logarithm of both sides of Eq. (2.4),[7] the result is

$$\log e_\lambda = \log (C\lambda^{-5}) - (c/\lambda T) \log e,$$

which can be rewritten as

$$\log e_\lambda = \log (C\lambda^{-5}) - (c \log e/\lambda)(1/T).$$

This has the form $y = a + bx$, where $y = \log e_\lambda$, $a = \log (C\lambda^{-5})$, $b = -(c \log e/\lambda)$, and $x = 1/T$. Thus Wien's formula implies that when $\log e_\lambda$ for a fixed λ is plotted against $1/T$, the resulting curve, called an isochromat, should be a straight line; the slope of the line is proportional to c, and the intercept on the $\log e_\lambda$ axis can be used to compute C. The value of C might vary from one series of observations to another, but it should be constant throughout any one series; the values obtained for c, on the other hand, should all be nearly the same.

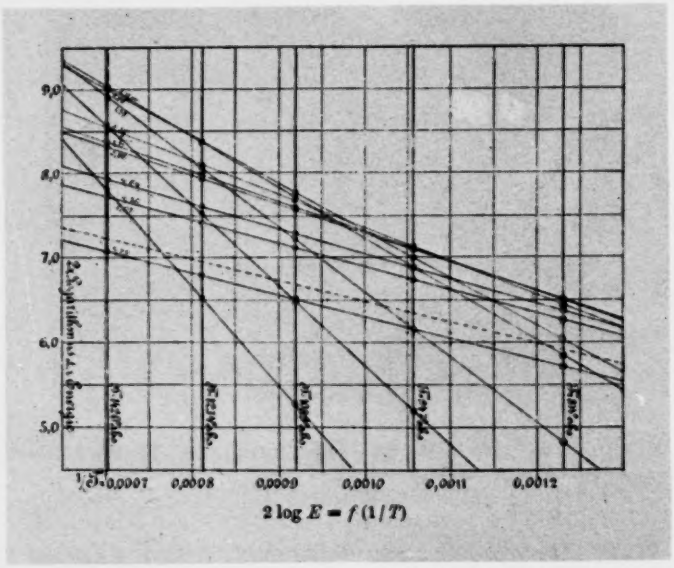

Fig. 2.3 A set of isochromats from Lummer and Pringsheim's first report on black body radiation.

In their earliest report, Lummer and Pringsheim found that the isochromats seemed indeed to be straight lines (see Fig. 2.3), but gave values of C and c that varied with wavelength. At that stage, they were not sufficiently confident in their procedures to regard the question as settled. By the time of the second report, they were soundly enough convinced of the invalidity of Eq. (2.4) that they looked for, and reported finding, curvature in the isochromats. They say, however, "Nevertheless, before we pass final judgment against the validity of the Wien-Planck equation [Planck had supported Wien's deduction on the basis of a different line of argument; see below], we consider it necessary to extend the studies over a larger temperature interval and wavelength range."[8] Finally, in the third report, the evidence had become unquestionable; the curvature in the isochromats was obvious (see Fig. 2.4 on next page). They firmly conclude that "Likewise brought down to the ground therewith are all those extensive consequences

[6] They state that this may have been the result of a poor adjustment which allowed the spectrometer to "look at" part of the cooler outer surface of the cavity as well as at the interior.

[7] The reader is reminded that the logarithm of a number y is the number a such that $10^a = y$; that therefore $10^{\log y} = y$; that if two numbers are equal, so are their logarithms; that $\log (xy) = \log x + \log y$; and that $\log (x^s) = \log [(10^{\log x})^s] = \log (10^{s \log x}) = s \log x$.

[8] Translated by G. L. T.

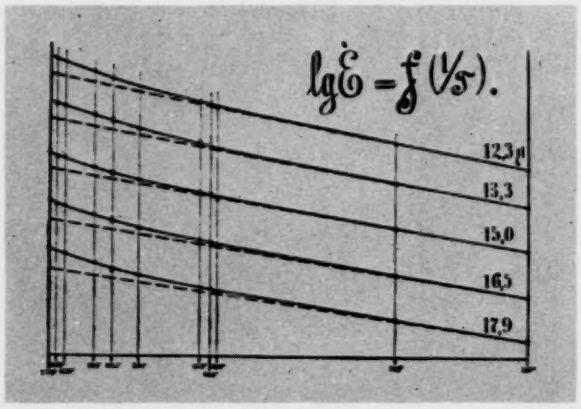

Fig. 2.4 A set of isochromats from Lummer and Pringsheim's third report.

that people have derived from the Wien-Planck equation."[8]

A discussion of Max Planck's introduction of the constant (customarily designated by h), that bears his name, while not exactly part of the account of the experiment, is relevant and interesting. Planck had for some time been interested in the problem of blackbody radiation, being attracted to it by the "absolute" character of the distribution law - its independence of the material of the cavity walls. He made use of this independence in his work on the problem by taking as the walls an assemblage of harmonic oscillators. His approach was always through the thermodynamics of the assemblage, with particular emphasis on the thermodynamic quantity called entropy. This quantity is a measure of the disorder in a system, and the second law of thermodynamics, with which Planck was thoroughly familiar, states that the natural tendency of a system is to change in such a way that its entropy increases. The equilibrium state, consequently, is the state of maximum entropy; and for a cavity, the equilibrium state is characterized by the cavity being filled with blackbody radiation. Planck's task, therefore, was to calculate the entropy of his assemblage.

In his earlier work, Planck was not familiar with the "disorder" interpretation of entropy, and he considered the entropy of an individual oscillator, which he sought to relate to its energy U. He found that a basic quantity was the curvature R of the graph of this relationship, and an erroneous assumption led him to conclude that R must depend on U through the equation

$$R = -\alpha/U, \qquad (2.5)$$

where α is a positive quantity that might depend on frequency. The radiation law to which this led was just Eq. (2.4), with $\alpha = \lambda/c$.

The correct radiation law, in fact, was first obtained by a purely empirical procedure. The experimental results showed that Eq. (2.5) needed modification for large values of U. If we write the equation as $1/R = -U/\alpha$, it can be seen that the simplest alteration of the sort needed would be to add a term in U^2. Planck did so, and obtained a radiation law of the form

$$e_\lambda \propto \lambda^{-5}/(e^{b/\lambda T}-1) \qquad (2.6)$$

which proved to fit the experimental results extremely well.

It was only after he had come to a more complete understanding of entropy that Planck was able to justify Eq. (2.6). The properties of entropy show that it is a measure of the probability of the state involved. The probability, in turn, can be found simply by counting the number of different microscopic arrangements - in the present case, the number of ways of assigning energies to the individual oscillators - compatible with the given over-all state, and by assuming that each microscopic arrangement is equally probable. In order for the counting process to be possible, however, the energy cannot be a continuous variable but must be parceled out in multiples of a basic unit ϵ, so that $U = n\epsilon$. When Planck put these concepts into his treatment, he found that the entropy S depended on U and ϵ only through the combination U/ϵ. Wien's displacement laws, on the other hand, implied that $S = f(U/\nu)$, where

ν is the frequency of the oscillator. Consequently, one must have $\epsilon = h\nu$: The energy of an oscillator must be an integral multiple of a basic unit proportional to the frequency.

The resulting radiation law was

$$e_\lambda = \frac{8\pi c^2 h}{\lambda^5 (e^{ch/k\lambda T} - 1)},$$

where h is Planck's constant and k is another universal constant, and c is the speed of light; this reduces to Wien's law for small values of λT, and to Rayleigh's law for large values of λT.

The beginnings of quantum theory, then, lay in an experiment whose results could be understood only by the introduction of an idea foreign to classical theory: that in some systems, energy is not infinitely subdivisible, but is exchanged with the rest of the universe only in discrete amounts, or quanta.

3 COLLISIONS OF ELECTRONS WITH ATOMS

In the early years of the twentieth century, the nature of the atom was very unclear. Obviously, the atom contained electrons, whose motions were related in some fashion to the frequencies of light emitted by the atom; but how the electrons and the remaining parts of the atom were fitted together remained a mystery. Meanwhile, the atom itself provided an object of study, and its properties were studied with interest.

One property is the ionization potential: the amount of energy that must be supplied to knock one electron loose from the atomic structure. One way to supply the energy is by striking the atom with an electron, and the energy is measured in terms of the electrical potential difference through which the electron must be accelerated to produce ionization. This potential difference is called the ionization potential.

J. Franck and G. Hertz, working at the University of Berlin, had supposedly measured the ionization potentials of several substances that are gases at ordinary temperatures. The method was to produce electrons by means of a heated filament; accelerate them through a measured, variable potential difference V maintained between the filament and a grid; and decelerate them again by a potential difference, which was ten volts greater than V, between the grid and a collector plate. In the space between grid and collector, the electrons could collide with gas atoms. They could not under any circumstances reach the collector. However (the argument ran), if V were greater than the ionization potential, the electrons would ionize some atoms by collision; the electric field would accelerate the positive ions toward the collector, which would then register a current. The value of the ionization potential, then, was the value of V for which current began to flow to the collector. By this method, Franck and Hertz measured what appeared to be ionization potentials for a half-dozen gases.

They hoped to correlate the ionization potentials with atomic radii, and for this purpose they wanted to make measurements on metallic atoms. They felt that maintaining their apparatus at the higher temperatures needed to produce metal vapors would produce errors - presumably, the currents they worked with were so small that the decrease in the resistance of the glass envelope, produced by the increase in temperature, would give rise to stray currents large enough to mask the desired effect. Accordingly, they devised a new form of apparatus. The grid N, instead of being fairly close (5 mm) to the filament, was made about 4 cm away; the collector plate G was placed only 1 or 2 mm from the grid, instead of some 2.5 cm. The potential between filament and grid was variable and such as to accelerate the electrons, as before; the potential between grid and collector was again decelerating, but small and constant.

The principle of the method is based on some assumptions which warrant brief discussion. The first is that when an electron and an atom undergo an elastic collision,[9] the electron loses only a negligible amount of energy. This assumption can be checked merely by use of the laws of conservation of energy and momentum,

[9]The reader is reminded that collisions are classed as elastic or inelastic according to what becomes of the initial kinetic energy. If it remains as kinetic energy of the two colliding bodies, though perhaps shared differently, the collision is elastic. If some of it is absorbed so as to alter the internal state of one of the colliding bodies, then the collision is inelastic.

and turns out to be completely valid.[10] The second is that a collision between an electron and an atom will be elastic as long as the kinetic energy of the electron is less than the ionization energy. This will be discussed later; but it seemed reasonable in the light of the earlier experiments. The third assumption is that the probability of a collision being inelastic, if the electron energy is large enough for that to be possible, is not very small compared to unity. This assumption, also, seemed to be borne out by the earlier work.

The actual operation was well described by Franck and Hertz in the report on their work, published in the Verhandlungen der Deutschen physikalischen Gesellschaft in 1914:

> As long as the accelerating potential is less than the decelerating, the current [to the collector] is null. Then it will increase, until the accelerating potential has become equal to the ionization potential. At this point, the electrons will undergo inelastic collisions in the neighborhood of the grid and thereby ionize. Since they themselves and the electrons set free by ionization traverse only a very small potential until their passage through the grid, they pass through it without an appreciable velocity and are incapable of running against the retarding field. The galvanometer current will, therefore, drop to zero as soon as the accelerating potential has become greater than the ionization potential. If one increases the accelerating potential further, the point at which the electrons undergo inelastic collisions moves inward from the grid. The electrons present after the inelastic collisions, therefore, pass on the way to the grid through a potential which is equal to the difference between the accelerating potential and the ionization potential. As soon as this difference has become greater than the constant retarding potential between N and G, electrons can again go against the retarding field and the galvanometer current again grows. Since the number of electrons is increased by the ionization, it actually grows larger than the first time. However, as the accelerating potential becomes equal to twice the ionization potential, the electrons undergo inelastic collisions a second time in the neighborhood of the grid. Since they thereby lose all their energy, and the newly produced electrons likewise have no appreciable velocity, no more electrons can run against the retarding field. Thus, as soon as the accelerating potential is greater than twice the ionization potential, the galvanometer current again drops to zero. Since the same phenomenon repeats itself each time the accelerating potential is equal to an integral multiple of the ionization potential, we should expect a curve having maxima of increasing size, whose separation is equal to the ionization potential.[11]

Just such curves were obtained; an example is reproduced in Fig. 3.1 (see next page). The maxima turned out to be quite sharp - the report gives a feeling that they were sharper than the authors had dared to expect - and Franck and Hertz imply a confidence that their results are accurate to 0.1 volt as compared with the 1-volt accuracy of the older method. The value obtained for the ionization potential of mercury was 4.9 volts. As a comparison with the older method, they remeasured the value for helium and found that the two methods gave very satisfactory agreement.

[10] If a body of mass M, initially at rest, is struck by a body of mass m, and if the collision is elastic, the incident body loses a fraction of its energy which is at largest $4mM/(M + m)^2$.

[11] Translated by G. L. T.

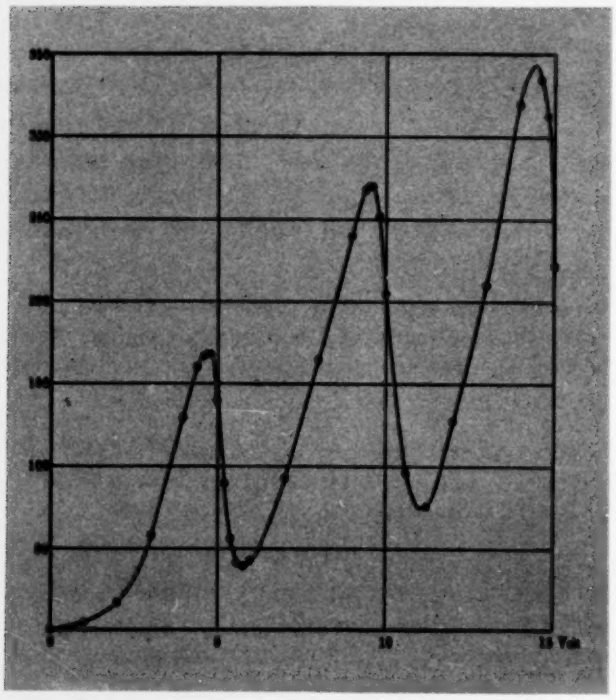

Fig. 3.1 Plot of collector current versus accelerating potential, showing the equal spacing of the maxima. (Note: Several effects combine to cause the spacing between maxima to be slightly different from the position of the first maximum.)

Such accuracy exceeded anything previously obtained, and enabled Franck and Hertz to make a quantitative test of a theoretical proposal that had been put forth several times: that the ionization energy should be equal to Planck's constant h times the frequency of one of the "proper motions" of the electrons. They felt it natural to choose a frequency that was very strongly absorbed by mercury vapor, that corresponding to a wavelength of 2536 Å. The potential indicated by the theory was 4.84 volts, in excellent agreement with the measured value.

At this stage, perhaps out of the very excellence of the agreement, doubts began to arise. The possibility presented itself that the electrons lost their energy not in ionizing the mercury - after all, in this method ionization had not been directly observed - but in exciting radiation. They could not check this new possibility with the same apparatus, as the wavelength 2536 Å lies well into the ultraviolet, while the glass envelope of their tube was opaque to ultraviolet light. Consequently, they built a new tube out of quartz, which is transparent to ultraviolet light. This tube was much simpler than the other, involving merely a platinum filament to provide electrons, a platinum grid toward which the electrons could be accelerated and on which they were collected, and a pool of mercury to provide vapor. The whole bulb was heated to about 150°C by means of a gas burner, and any radiation given off was analyzed by means of an ultraviolet spectrograph.

The results were somewhat surprising. When the potential between filament and grid was less than 4.9 volts, no radiation was emitted by the mercury vapor.[12] When the potential was greater than 4.9 volts, the mercury radiated, as expected; but it radiated only the wavelength 2536 Å, despite the fact that many lines of the mercury arc spectrum are more intense than that at 2536 Å and despite the fact that the wavelengths of many of these lines correspond to potentials less than 4.9 volts. Franck and Hertz concluded that in some collisions, the energy that the electron had acquired, if large enough, was converted into radiation; they remained convinced that in other collisions at the same energy, the energy was used to ionize the atom.

We should now abandon the point of view of Franck and Hertz, and examine these conclusions in the light of later knowledge. In fact, if Franck and Hertz had been able to work with mercury or an alkali metal in both the earlier apparatus and the later one, they would have met a peculiar inconsistency: For any of these substances, the two methods would have given different results, although they gave identical results for helium and would have done so for any of the substances

[12] The critical value in this arrangement actually turns out to be somewhat less than 4.9 volts because the electrons already have some energy when they are produced by the filament.

actually measured in the earlier apparatus. The reason is that in no case was there actually an ionization potential measured. The increases in collected current which Franck and Hertz ascribed to positive ions had a much different origin. The atoms or molecules were being excited and were emitting radiation, and the radiation was producing photoelectrons from the collector plate. The radiation emitted by mercury, and the radiation that would have been emitted by an alkali metal, could not have produced this effect.[13] In every case, the ionization potential is higher than the value measured by Franck and Hertz in this work.

What, then do the measured values mean? They certainly are thresholds of energy beyond which <u>something</u> happens to the <u>internal</u> behavior of the atom. When the current to the collector drops, it is a signal that the electrons have gained just enough energy to lose most of it in an inelastic collision with an atom. When it drops a second time, the electrons are gaining enough energy to experience such an encounter twice during their travel from filament to grid, and so on. Evidently there is a minimum amount by which the internal energy of the atom can be changed - and in that statement, forced on us by the experimental facts, is found the failure of extrapolation from experience on the ordinary scale. On the ordinary scale, it is perfectly acceptable for a dynamical system to have characteristic <u>frequencies</u>, as an atom does; a pendulum or a stretched string comes immediately to mind, and other more complicated examples could be cited. But the <u>energy</u> of any such system can be changed by an arbitrary amount. Apparently, the same is <u>not</u> true of an atomic system. It can exist only in certain states with certain, distinct energy values. Its state can change only from one of these to another, and so its energy can change only by certain distinct amounts. This is the lesson of the Franck-Hertz experiment. And even as the experiments were being done, Niels Bohr was taking it as one of the basic postulates on which to build his theory of the atom, quite unaware that his radical proposal was being tested and was passing the test.

[13] For a further discussion relating to this point, see Chapter 4.

4 THE PHOTOELECTRIC EFFECT

We have seen in Chapter 2 that the quantum concept was originally introduced as an aspect of the behavior of radiating oscillators rather than of the radiation itself. It was natural, however, to feel that such behavior might impress itself at least partially on the radiation. The quantum concept had existed only five years when Albert Einstein seized upon this possibility as an explanation for the peculiar effects of radiation in the photoelectric effect. It was 1914 before a thorough, convincing test of Einstein's ideas was reported.

The photoelectric effect was discovered late in the nineteenth century, and by 1914 had been studied fairly extensively. The basic phenomena were known: A beam of light striking the surface of a metal liberates electrons from the metal, provided its frequency is greater than a critical value dependent on the kind of metal. The electrons acquire in the process some kinetic energy, the amount of which increases with increasing frequency of the light according to a relationship whose form had not been experimentally established with certainty in 1914. If the electrons are collected at another electrode and made to constitute a current, the magnitude of the current is proportional to the intensity of the stimulating light. The entire process is virtually instantaneous.

In a sense, the very existence of the photoelectric effect can be regarded as a failure of extrapolation, as there was nothing in classical theory that would have suggested such a process. Even granted its existence, however, all but one of the properties just listed are in conflict with what would be expected from classical theory. The dependence of photocurrent on light intensity is quite reasonable. But any "reasonable" assumption about the mechanism would involve an interaction of the electron with the electric field of the light wave, and the intensity of an electric field wave is independent of the frequency. The appearance of the frequency as an essential factor in the phenomenon, therefore, is quite unexpected. The time dependence is an even more drastically shocking result. The most favorable assumption regarding the transfer of energy from the light to the electron is that some sort of resonance process takes place; this assumption implies that the electron will absorb all the energy incident on an area one wavelength square. Computations on this basis lead to the conclusion that for a beam of very low intensity, but still sufficient for easy observation of the photoelectric effect, an electron would take about 500 years to accumulate energy equal to that observed.

Einstein's proposal of an alternative theory had led R. A. Millikan to carry out an exhaustive experimental study at the Ryerson Laboratory of the University of Chicago. His results were first reported to a meeting of The American Physical Society in April, 1914; and a detailed description was published in The Physical Review in 1916. The status of the new theory is well described in the following quotation from the introduction to that paper:

> It was in 1905 that Einstein made the first coupling of photo effects and with [sic] any form of quantum theory by bringing forward the bold, not to say the reckless, hypothesis of an electro-magnetic light corpuscle of energy $h\nu$, which energy was transferred upon absorption to an electron. This hypothesis may well be called reckless first because an electromagnetic disturbance which remains localized

in space seems a violation of the very conception of an electromagnetic disturbance, and second because it flies in the face of the thoroughly established facts of interference. The hypothesis was apparently made solely because it furnished a ready explanation of one of the most remarkable facts brought to light by recent investigations, viz., that the energy with which an electron is thrown out of a metal by ultraviolet light or x-rays is independent of the intensity of the light while it depends on its frequency. This fact alone seems to demand some modification of classical theory or, at any rate, it has not yet been interpreted satisfactorily in terms of classical theory.

While this was the main if not the only basis of Einstein's assumption, this assumption enabled him at once to predict that the maximum energy of emission of corpuscles under the influence of light would be governed by the equation

$$\tfrac{1}{2}mv^2 = V \cdot e = h\nu - p, \qquad (I)$$

in which $h\nu$ is the energy absorbed by the electron from the light wave, which according to Planck contained just the energy $h\nu$, p is the work necessary to get the electron out of the metal and $\tfrac{1}{2}mv^2$ is the energy with which it leaves the surface, an energy evidently measured by the product of its charge e by the P.D. against which it is just able to drive itself before being brought to rest.

At the time at which it was made this prediction was as bold as the hypothesis which suggested it, for at that time there were available no experiments whatever for determining anything about how P.D. varies with ν, or whether the hypothetical h of equation (I) was anything more than a number of the same general magnitude as Planck's h. Nevertheless, the following results seem to show that at least five of the experimentally verifiable relationships which are actually contained in equation (I) are rigorously correct. These relationships are embodied in the following assertions:

1. That there exists for each exciting frequency ν, above a certain critical value, a definitely determinable maximum velocity of emission of corpuscles.
2. That there is a linear relation between V and ν.
3. That $dV/d\nu$ or the slope of the $V\nu$ line is numerically equal to h/e.
4. That at the critical frequency ν_0 at which $\nu = 0$, $p = h\nu_0$, i.e., that the intercept of the $V\nu$ line on the ν axis is the lowest frequency at which the metal in question can be photoelectrically active.
5. That the contact E.M.F.[14] between any two conductors is given by the equation

$$\text{Contact E.M.F.} = h/e(\nu_0 - \nu_0') - (V_0 - V_0').$$

No one of these points except the first had been tested even roughly when Einstein made his prediction and the correctness of this one has recently been vigorously denied by Ramsauer. As regards the fourth Elster and Geitel had indeed concluded as early as 1891, from a study of the alkali metals, that the more electro-positive the metal the smaller is the value of ν at which it becomes photo-sensitive, a conclusion however which later re-

[14] When two different metals are electrically connected, either by direct contact or through an electrical circuit, a difference in electrical potential exists between them. This difference is the contact potential difference or contact emf. It does not affect the flow of current in the circuit, as the sum of all the contact emf's around a complete circuit vanishes. It does, however, contribute to static effects.

searches on the non-alkaline metals seemed for years to contradict.

During the ten years which have elapsed since Einstein set up his equation the fifth of the above assertions has never been tested at all, while the third and fourth have never been subjected to careful experimental test under conditions which were even claimed to permit of an exact and definite answer, nor indeed can they be so subjected without simultaneous measurements in vacuo of both contact potentials and photo-potentials in the case of metals which are sensitive throughout a long range of observable frequencies. In making this statement I am not underrating at all the exceptionally fine work of Richardson and Compton, who in common with most other observers interpreted their results in terms of Einstein's equation, but who saw the significance of that equation much more clearly than most of their predecessors had done. I am merely calling attention to the fact that the slope mentioned in (3) and the intercept mentioned in (4) cannot possibly be determined with any approach to certainty unless the region of wave-lengths open to study is larger than it is in the case of any save the alkali metals, and also, in the case of (4) unless simultaneous measurements are made in vacuo upon photopotentials and contact E.M.F.'s.

As the last paragraph of the foregoing quotation implies, there had been earlier attempts at verification of Einstein's proposal. In particular, very careful studies had been reported in 1912 by A. Ll. Hughes from Cambridge, England, and, as mentioned, by O. W. Richardson and A. H. Compton from Princeton. Nevertheless, a review of the subject published in 1913 concluded that the case was still open.

The work at the University of Chicago had begun in 1905, presumably immediately upon publication of Einstein's paper, and various aspects of the work had been reported in the years from 1907 to 1912. Gradually, these earlier results, as Millikan puts it,

> ...revealed the necessity of questioning the validity of all results on photopotentials unless the effects of surface films are eliminated either by removal of the films or by simultaneous measurement in vacuo of photopotentials and contact E.M.F.'s or by both procedures at once. Accordingly I have initiated in 1910 on a somewhat elaborate scale simultaneous measurements on photoeffects and contact E.M.F.'s in vacuo on film free surfaces.

The investigation required a number of more or less elaborate precautions, some of which will be discussed later. The most bothersome, however, was the one mentioned in the last quotation, that of eliminating the effects of surface films. The obvious solution is to prepare the surfaces under such conditions that the films are prevented from forming; and, despite the formidable mechanical difficulties involved, this was the course taken.

> In all of this photoelectric work at the Ryerson Laboratory the same general method has been employed, namely, the substances to be studied or manipulated have been placed in the best obtainable vacuum on an electromagnetically controlled wheel and all the needed operations have been performed by movable electromagnets placed outside. As new operations have been called for the tubes have by degrees become more and more complicated until it has become not inappropriate to describe the present experimental arrangement as a machine shop in vacuo. The operations which are now needed in all the tubes which are being used are:
>
> 1. The removal in vacuo of all surface films from all surfaces.

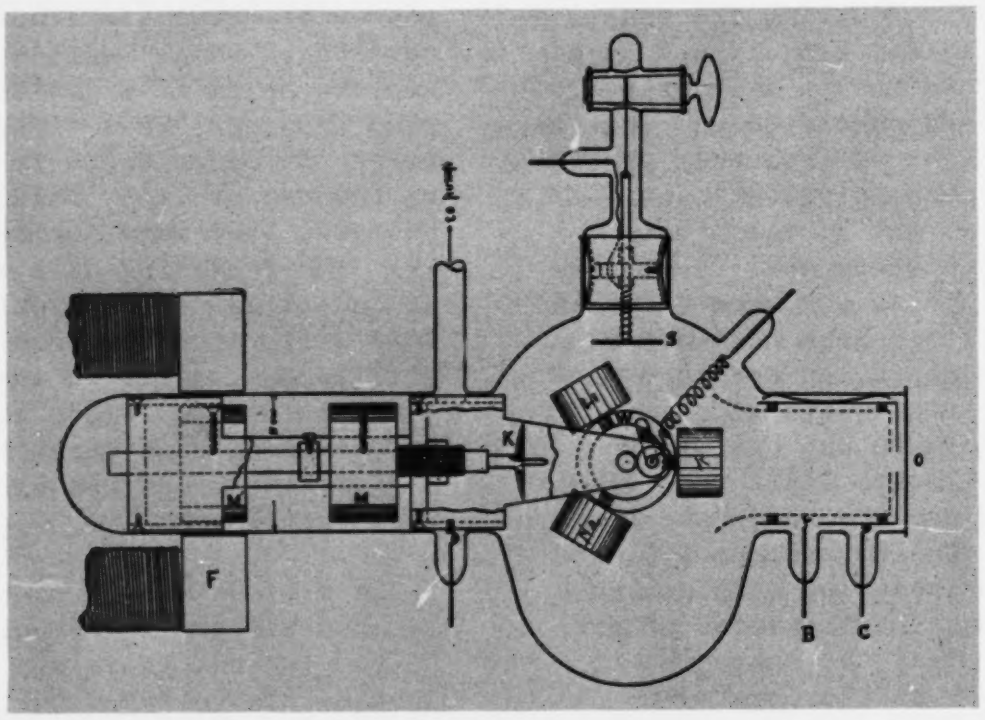

Fig. 4.1 Diagram of one of the tubes used by Millikan in his study of the photoelectric effect.

2. The measurement of the photocurrents and photopotentials due to these film-free surfaces.
3. The simultaneous measurement of the contact E.M.F.'s of the surfaces.

A diagram of one of the tubes appears in Fig. 4.1. The three cylinders carried on the wheel W are cast from the metals to be studied. The wheel itself can be rotated by an electromagnet, not shown, so as to bring any of the cylinders opposite any of the other parts of the apparatus. K is a rotating knife, which can be moved back and forth along the axis, and rotated, by the action of the electromagnet F on the armatures M and M'. The metal to be studied is first brought opposite the knife, which is then advanced far enough to cut a thin shaving off the face of the cylinder; the cut is made by rotating the knife, which is then retracted again, the shaving falling down below the wheel where it helps to remove any residual oxygen from the bulb. The fresh surface can then be turned opposite either the plate S, for measurement of the contact emf, or the window O and electrodes B and C, for measurement of photoelectric behavior.

Electrodes S and B were made of copper, carefully treated so as to have identical contact potentials. The measurement of the contact emf was based on the fact that if there were a potential difference, of whatever origin, between S and the cylinder, a change in the separation of S and the cylinder would cause some charge to move through an electrometer connected to them. Thus, when an external potential was applied so as to give no motion of charge, the external potential must be just canceling the contact potential.

In studies of the photocurrents themselves, a beam of monochromatic light entered through the window O and struck the surface being studied. The photoelectrons were collected by the double cylinder B and C, which were insulated from each other inside the tube but, in the actual measurements, electrically connected outside.

One precaution is discussed in the following quotation from the paper:

> Since the aim was to test with the utmost possible accuracy the slope

> of the line connecting frequency with the maximum P.D. it was necessary first to know ν with great precision and second to see that no trace of light of frequency greater than that being plotted[15] got through the slit of the spectroscope. To this end a . . . mercury lamp was used as a source . . . and only such lines were chosen for use as had no companions anywhere on the short wave-length side.... Light filters to cut out stray short wave-length light were also used.... Since the measurement was to be made on the maximum P.D. and since this increases with decreasing wave-length it was not of great importance that the source be of great purity on the long wave-length side.

Millikan also considered and, as far as possible, eliminated errors which had plagued other workers. One was "back leakage" of photoelectrons liberated from the collecting electrode by reflected light. If such electrons are present, the apparent critical retarding potential will actually be the potential at which the "forward" current just balances the back leakage. Richardson and Compton had been particularly conscious of this problem. Millikan avoided it, for all but one of the wavelengths he used, by using a collecting electrode whose own photoelectric threshold wavelength was shorter than that of the incident light. Another source of error was the fact that previous investigators had used a very small range of wavelengths, none extending over a range equal to the smaller limit, so that the workers were forced to try to deduce the shape of a curve from a very short portion of it. Millikan used a range nearly four times the lower limit. A third source of error was light of short wavelength which reached the sensitive surface by diffuse reflection in the monochromator. This problem was eliminated, in cases where it was critical, by the use of light filters; usually, however, they were not necessary.

The resulting data were plotted as curves of photocurrent versus potential difference for each of several wavelengths. A set of such curves is shown in Fig. 4.2 (see next page). The intercept of each curve with the horizontal axis gives the value of V for that wavelength.

> It will be seen...that the maximum possible error in locating any of the intercepts is say two hundredths of a volt and that the total range of volts covered by the intercepts is more than 2.5. Each point, therefore, of a potential-frequency curve should be located with not more than a per cent of uncertainty. The frequencies are, of course, known with great precision.

The values of the intercepts were then plotted versus wavelength. The plot obtained from the curves of Fig. 4.2 is shown in Fig. 4.3 (see next page). "It will be seen that the first result is to strikingly confirm the conclusion...as to the correctness of the predicted linear relationship between maximum P.D. and ν, no point missing the line by more than about a hundredth of a volt." If Eq. (4.1) is divided by e, the result is

$$V = (h/e)\nu - (h/e)\nu_0,$$

a straight line whose slope is h/e. Thus, from the measured slope of the line in Fig. 4.3 and Millikan's own earlier determination of e, he could compute h. The result was 6.56×10^{-27}, in complete agreement with the value originally computed by Planck from the constants of the blackbody radiation laws. Moreover, many determinations, on different surfaces, gave the same

[15] The reader should recall that, as stated in the first paragraph of this chapter, it was already known that an increase in frequency produced an increase in the energy of the emitted electrons.

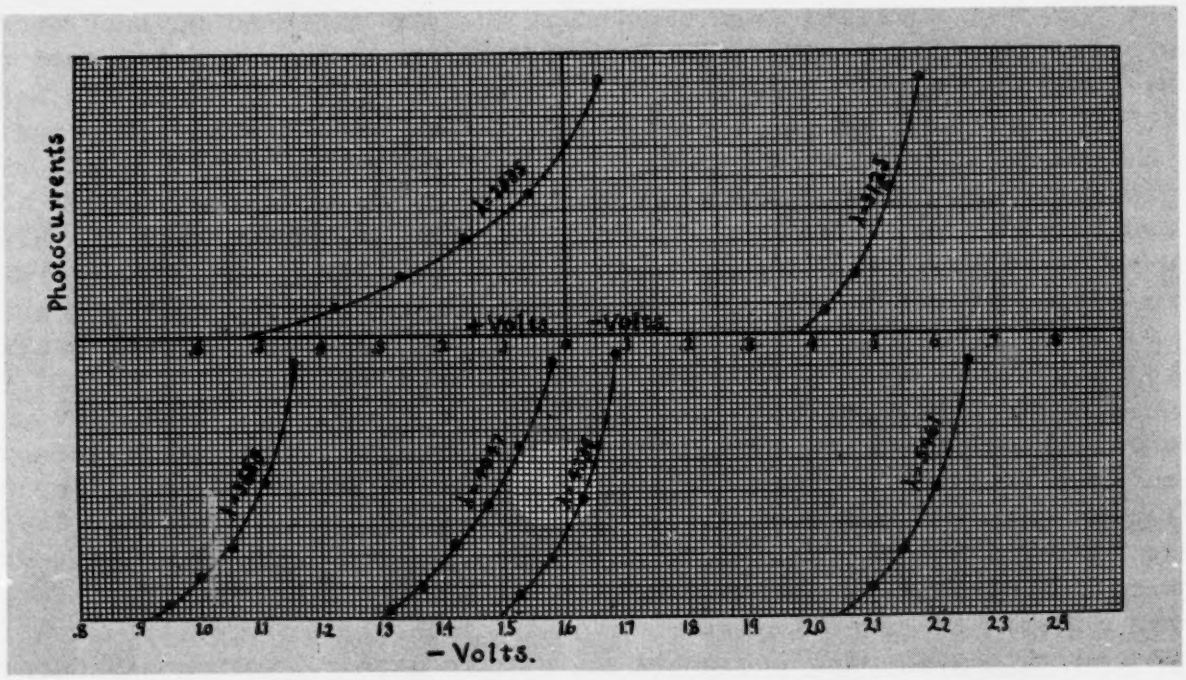

Fig. 4.2 A set of photocurrent curves for sodium. These curves were used to provide data for the curve in Fig. 4.3.

results to within experimental uncertainty.

Critical frequencies were determined in two ways. It will be recalled that the experimental apparatus was so devised that Millikan could measure the contact emf between the photosensitive surface and a surface identical to that of the collecting electrode. The actual potential difference between the two surfaces was equal to the measured, applied voltage minus the contact potential. The curve of actual potential difference versus frequency,

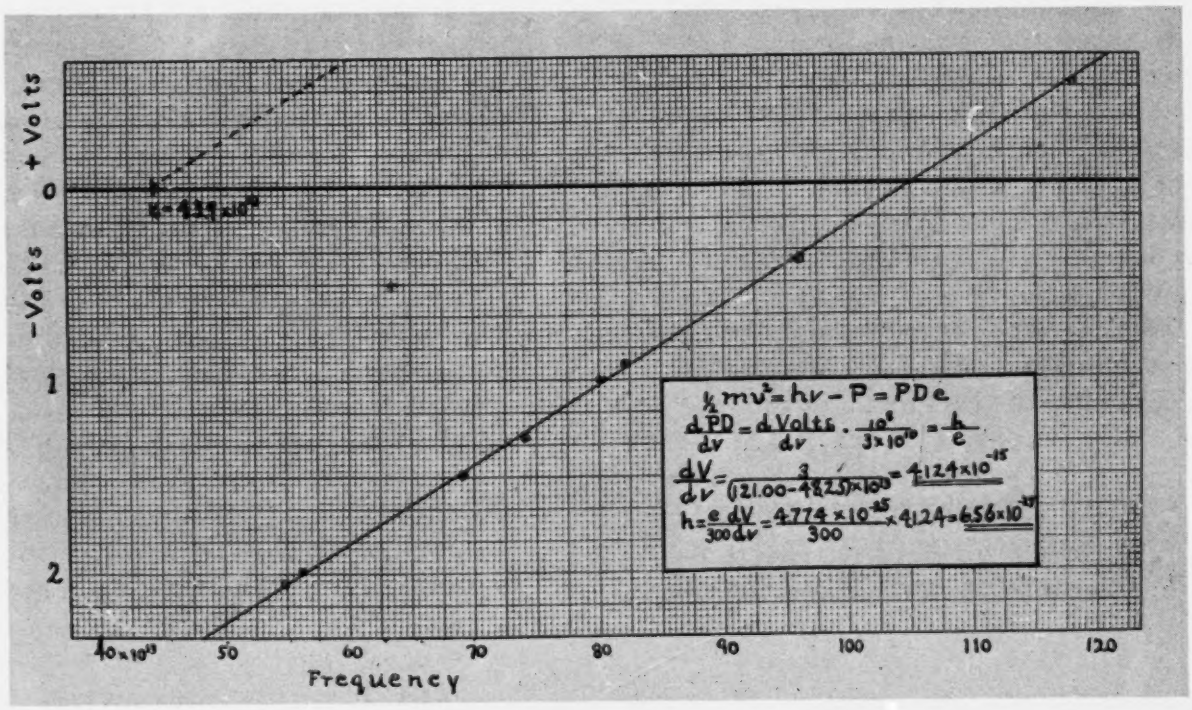

Fig. 4.3 The plot of maximum energy, measured as stopping potential difference, versus frequency for sodium. The inset outlines the computation of Planck's constant h from the measured slope of the line. The dashed curve is discussed in the text.

therefore, could be obtained from the "raw" curve of Fig. 4.3 by shifting it upward by an amount equal to the contact emf. The resulting curve is shown dashed in Fig. 4.3. It is this curve whose intercept with the frequency axis gives ν_0.

The second method made use of the fact that if the intensity and frequency of the stimulating light are fixed, and the potential difference used to accelerate the electrons toward the collector is varied, the collected current at first increases with increasing voltage but eventually reaches a limiting value, the saturation current. For fixed frequency, the saturation current is directly proportional to the intensity. The procedure was to measure the intensity of the incident radiation, by means of a thermopile,[16] together with the corresponding saturation photocurrent, for each of several wavelengths. The saturation photocurrent per unit intensity was then plotted against wavelength. The point at which the curve crosses the wavelength axis is the critical wavelength λ_0, related to ν_0 by $\lambda_0\nu_0 = c$. This method might be expected to be less reliable than the first. Nothing was known in advance as to the likely shape of the curve; and, in the examples shown in Millikan's paper, the curve was determined only by three or four points and then extrapolated. Nevertheless, Millikan was confident that he could locate the intercept to within 100 Å; and certainly the results seemed to bear him out, at least as judged by agreement between the two methods.

[16] A thermopile is a group of thermocouples, connected electrically so as to act in combination and blackened so as to absorb incident radiation.

The fifth point mentioned in Millikan's introduction need not concern us further, except to mention that it was satisfactorily validated.

Millikan concludes his paper with five pages of discussion of theories of photoemission. Most of this material has been made obsolete by later developments. Two portions, however, remain valid and warrant quotation:

> Perhaps it is still too early to assert with absolute confidence the general and exact validity of the Einstein equation. Nevertheless, it must be admitted that the present experiments constitute very much better justification for such an assertion than has heretofore been found, and if that equation be of general validity, then it must certainly be regarded as one of the most fundamental and far reaching of the equations of physics; for it must govern the transformation of all short-wave-length electromagnetic energy into heat energy. ...
>
> The photoelectric effect then, however it is interpreted, if only it is correctly described by Einstein's equation, furnishes a proof which is quite independent of the facts of black-body radiation of the correctness of the fundamental assumption of the quantum theory, namely, the assumption of a discontinuous or explosive emission of the energy absorbed by the electronic constituents of atoms from ether waves. It materializes, so to speak, the quantity "h" discovered by Planck through the study of black body radiation and gives us a confidence inspired by no other type of phenomenon that the primary physical conception underlying Planck's work corresponds to reality.

5 MAGNETIC PROPERTIES OF ATOMS

Bohr's theory of atomic structure was published in a series of papers between 1913 and 1915,[17] and by the early 1920's had been fairly thoroughly developed by many workers, notably Arnold Sommerfeld. One of its fundamental postulates was that certain dynamical quantities relating to periodic motion could not take on arbitrary values, but only a set of discrete values, integral multiples of Planck's constant h. This postulate accounted nicely for the discrete energy states deduced from the experiment of Franck and Hertz described in Chapter 3 - not only for their qualitative existence, but, in the case of the hydrogen atom, for their quantitative values as well.

An even more curious result of the application of this rule is that if an atom whose complement of electrons has a nonzero angular momentum is in a magnetic field, the angular momentum vector can make only certain well-defined, discrete angles with the direction of the field. The angles permitted by the theory are those such that the compoent of the angular momentum along the field is an integral multiple (positive, negative, but not zero) of $h/2\pi$ where h is Planck's constant.[18] The reason that this result was regarded as curious was just the sort of extrapolation discussed in Chapter 1, for classically the value of the component of a vector along any specified direction is restricted only by the condition that it cannot exceed the magnitude of the vector.

In a paper in the Zeitschrift für Physik in 1921, Otto Stern made the following statement: "Now, whether the quantum theoretical or the classical conception is valid can be distinguished by means of an experiment, completely simple in principle. One needs only to study the deflection that a beam of atoms undergoes in a suitably inhomogeneous magnetic field."[19] The theoretical basis for the statement is as follows: An electron moving in one of the orbits of Bohr's theory, with an angular momentum $\vec{L}$, has a magnetic moment $\vec{\mu}$ which is proportional to $\vec{L}$ and along the same line. (In an atom containing several electrons, both the angular momenta and the magnetic moments of the individual electrons add vectorially, so that the total magnetic moment is related in this same way to the total angular momentum.) When the atom is in a magnetic field $\vec{B}$, the action of the field on the magnetic moment produces a torque in a direction perpendicular to both $\vec{\mu}$ (and therefore $\vec{L}$) and $\vec{B}$, and of magnitude proportional to B and to L, and dependent on the angle between $\vec{B}$ and $\vec{L}$. The effect, according to classical dynamics, is to cause $\vec{L}$ to precess about $\vec{B}$ - to trace out a cone whose axis is along $\vec{B}$. The frequency of precession is proportional to B, the factor of proportionality involving only the properties of the electron and natural constants; for a field of 10^3 gauss, the period is 7×10^{-10} second. Now suppose that the field is not uniform, but varies from point to point. Denote by $\partial\vec{B}/\partial x$ the change in $\vec{B}$ per unit displacement in the x direction, and similarly for y and z. Then there will be a net force $\vec{F}$ on the atom, given by

[17] James Franck remarked in 1961 that he and Hertz had not known of Bohr's work while their own experiments (see Chapter 3) were going on, but that even if someone had told them it existed they probably would not have read it: The problem of atomic structure was regarded as so arcane that anyone claiming to have solved it was likely to be a crackpot.

[18] The theory also maintained that the magnitude of the angular momentum must be an integral multiple of $h/2\pi$.

[19] Translated by G. L. T.

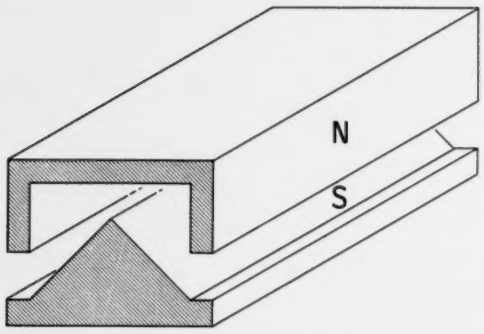

Fig. 5.1 Shape of magnet poles to give a large value of $\partial\vec{B}/\partial z$ in the direction of $\vec{B}$.

$$\vec{F} = \mu_x \partial\vec{B}/\partial x + \mu_y \partial\vec{B}/\partial y + \mu_z \partial\vec{B}/\partial z.$$

If this is averaged over a time long compared with the period of precession, the first two terms average to zero because μ_x and μ_y vary sinusoidally, and there remains only

$$\langle\vec{F}\rangle = \mu_z \partial\vec{B}/\partial z.$$

The atom, therefore, is accelerated in the direction of $\partial\vec{B}/\partial z$.

Now imagine an electromagnet whose poles are shaped as shown in Fig. 5.1. Close to the knife-edge, the field $\vec{B}$ is strongest and the direction of $\partial\vec{B}/\partial z$ is along $\vec{B}$. An atom of mass M traveling along parallel to the knife-edge and just above it will experience an acceleration $a = \langle F\rangle/M = \mu_z (\partial B/\partial z)/M$, in a direction along $\partial\vec{B}/\partial z$ and thus perpendicular to its initial motion. If it spends a time t traversing the field, it will have been deflected from its original path by an amount $x = \frac{1}{2}a\, t^2 = \frac{1}{2}\mu_z(\partial B/\partial z)t^2/M$. The time t is the speed v at which the atom travels divided into the length ℓ of the magnet poles. The deflection, then, is

$$x = \frac{1}{2}\,\frac{\mu_z}{M}\,\frac{\partial B}{\partial z}\,\frac{\ell^2}{v^2}. \qquad (5.1)$$

Now consider an atom for which $L = h/2\pi$. Then according to quantum theory, the component L_z of $\vec{L}$ along $\vec{B}$ can only be $\pm h/2\pi$, and μ_z can be only $\pm\mu$. An initial beam of atoms all having the same velocity would then be split into two by the action of the field, and there would be no part of the beam undeflected.

The classical case is a bit more complicated. Let us write Eq. (5.1) as

$$x = C\mu_z$$

or

$$x = C\mu \cos\theta, \qquad (5.2)$$

where θ is the angle between $\vec{\mu}$ and $\vec{B}$. The classical concept is that θ can have all possible values, so that $\cos\theta$ will range continuously from -1 to $+1$. The initial beam is not split but spread, with the extent of spread being equal to the amount by which the two components predicted by quantum mechanics would be separated. However, it is still possible that the classical result would look somewhat like the quantum one, if the intensity along the spread had maxima at the ends and a minimum at the center. We can show that this is not the case. The intensity distribution is the number of atoms per unit deflection; the procedure, then, is to find the number of atoms for which θ is in a small range $d\theta$ and the values of the deflection corresponding to this same range, and take the ratio. It is shown in calculus that if x and θ are related as in Eq. (5.2), then the range dx of x corresponding to a very small range $d\theta$ of θ is given by

$$dx = C \sin\theta\, d\theta, \qquad (5.3)$$

which is then the denominator of our ratio. The numerator, at least as regards its dependence on θ, is obtained by recognizing that the numbers for two different angles θ_1 and θ_2 are to each other as the areas of two narrow bands around a sphere, one at colatitude[20] θ_1 and one at colatitude θ_2, having equal widths. Figure 5.2 shows such a band. Its area is the circumference of the band, $2\pi a$, times the

[20] The colatitude of a point on a sphere is the angle that a radius to that point makes with the axis of the sphere. It is 90° minus the latitude.

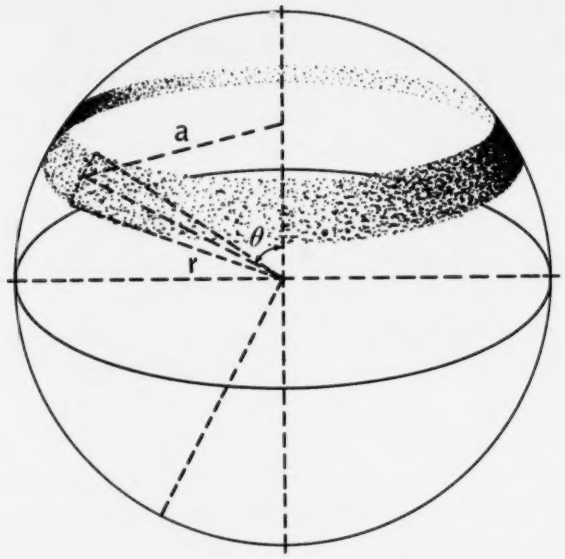

Fig. 5.2 Band at colatitude θ on a sphere of radius r.

width (r dθ), provided dθ is small. The radius a is equal to sin θ times the radius r of the sphere, so that the area of the band is $2\pi r^2 \sin\theta\, d\theta$. In the ratio of two such areas, all factors cancel out except $\sin\theta_1/\sin\theta_2$. Thus, the number of atoms whose magnetic moments make an angle with $\vec{B}$ in the range dθ around θ is proportional to sin θ dθ. When this is divided by Eq. (5.3), the angular dependence cancels out. The result is that according to classical theory, the initial spot is spread into a band of uniform intensity.

These calculations, however, have assumed a beam involving a single velocity. In actuality, the beam is obtained by vaporization and contains a distribution of velocities. This turns out to improve matters. In the quantum case, each of the two spots is spread out somewhat, but the central region is still a <u>minimum</u> [Eq. (5.1) shows that if μ_z cannot be zero, and $\partial B/\partial z \neq 0$, a zero deflection can result only from an infinite velocity.] In the classical case, on the other hand, the single uniform streak is replaced by a superposition of streaks whose lengths range from very small to very large; the combination produces a <u>maximum</u> of intensity at the center. It is evident that the distinction is, as Stern maintained, clearcut.

Moreover, the experiment seems simple in principle. In practice, it is another matter. To begin with, the experiment must be carried out in vacuum so that the beam will not be destroyed by scattering from gas molecules. This restricts the possible length ℓ of the poles to a few centimeters. It is possible to make B of the order of 10^3 gauss, and $\partial B/\partial z$ of the order of 10^4 gauss per centimeter. The value of μ is known. The combination Mv^2 is twice the kinetic energy of the atom; and the kinetic energy is determined by the vapor temperature, which may be of the order of 1000°K. When such numbers are put into the formula, the deflection that can be expected is found to be of the order of 0.01 mm. The experiment was possible, but it would evidently be a delicate one.

When Stern submitted the paper containing the foregoing analysis, in late August of 1921, he and a coworker, Walther Gerlach, were already occupied with carrying out the experiment at Frankfurt am Main, Germany; and by the middle of November they had preliminary results - too preliminary, however, to permit a decision on the main question. They proceeded to make some improvements on the apparatus (a footnote to the third paper comments that "It was possible for these to be worked out and tested by joint efforts during the Christmas vacation"[21]), and submitted firm conclusions on 1 March 1922.

The experimental arrangement warrants fairly detailed description. The substance used was silver, which was vaporized in an electrically heated oven and escaped through a circular opening of area 1 mm^2. At a distance of 2.5 cm from the oven[22] was a dia-

[21] Translated by G. L. T.

[22] This distance was increased from the 1 cm used in the preliminary work, so as to prevent the aperture from being plastered over either by molten silver splattered out of the oven or by too rapid incrustation by deposition from the beam.

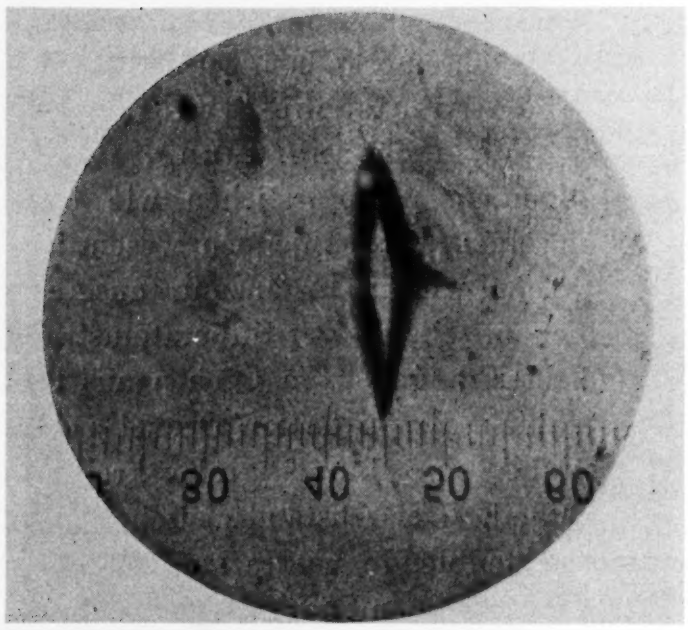

Fig. 5.3 Splitting of a beam of silver atoms by an inhomogeneous field. Each scale division is 1/20 mm.

phragm containing an approximately circular aperture of area 3×10^{-3} mm^2, i.e., radius about 0.03 mm. Another 3.3 cm beyond this was a second diaphragm, whose opening was in the form of a slit of length 0.8 mm and width 0.03 to 0.04 mm, oriented perpendicular to the direction of B. Such tiny dimensions were obviously necessary to produce a beam which was not vastly larger than the amount by which it was deflected. The slit was placed just at the apex of one end of the knife-edge pole piece (see Fig. 5.1), and the set of openings was so adjusted that the beam traveled parallel to the knife-edge. The magnet poles were 3.5 cm long. The whole arrangement was housed in a casing whose walls were thick enough to prevent pressure on the magnet poles from shifting the relative positions. The "exposure times" were eight hours; even then, the deposit on a glass plate at the far end of the pole pieces was too thin to be visible, and had to be darkened by precipitation of nascent silver.

The best exposure is reproduced in Fig. 5.3. Gerlach and Stern say:

> Two other exposures yielded a result identical in all essential points, not, however, with this complete symmetry. It must be said here that a reliable adjustment of such small diaphragms by optical means is very difficult, so that the achievement of such a completely symmetric exposure as in Fig. 3 is surely in part the result of luck; misplacement of one diaphragm by a few hundredths of a millimeter is already enough to make an exposure fail.[23]

The one characteristic common to all three exposures was the clear separation of the beam into two components. As the authors put it, "The splitting of the atomic beam in the magnetic field gives rise to two discrete beams. There are no undeflected atoms detectable."[23]

Reference to the preceding discussion shows that this result clearly confirms the quantum hypothesis as opposed to the classical behavior. The terminology of the time was used by Gerlach and Stern: "We behold in these experimental results the direct experimental proof of the quantization of direction in a magnetic field.[23]

[23] Translated by G. L. T.

6 THE SCATTERING OF X RAYS

X rays were identified as electromagnetic waves in 1912, by means of diffraction experiments suggested by von Laue and carried out by Friedrich and Knipping. The behavior of electromagnetic waves was well understood on the basis of Maxwell's theory. In particular, it was an easy matter to envisage a mechanism by which they were scattered, and it was a straightforward procedure to compute the quantitative features of the scattering. The mechanism was that the varying electric field of the wave set the electrons in the scatterer into forced vibration; the electrons, in turn, because they were being accelerated, emitted radiation. The scattered radiation was of the same frequency as the incident - it had to be, as it was just the frequency of vibration of the emitting electrons. It had all the properties of the radiation emitted by an oscillating electric dipole: intensity symmetrically distributed around the line of the electron's motion and, in any plane containing that line of motion, varying as the square of the sine of the angle between the line of motion and the direction of propagation; and polarization properties that are of no concern in the present discussion. Moreover, the fraction of the incident energy transferred into the scattered radiation should be independent of the frequency.

In the ten years following 1912, this theory met with steadily increasing difficulty. The first discrepancy was that for x rays of very short wavelength, or for γ rays, the scattered intensity was greater in the forward direction (relative to the incident radiation) than in the backward. This feature received, for a time, a quantitative explanation by ascribing to the electron a size comparable to the x-ray wavelength and allowing for interference between the rays scattered by different parts of the electron. As data accumulated, however, it was found that the value that must be assigned for the diameter of the electron varied with the wavelength of the incident radiation - obviously a most unsatisfactory situation. In addition, a still more serious difficulty had appeared. It was discovered that the scattered radiation was different in frequency from the incident. Arthur H. Compton, in a paper in The Physical Review in 1923, had this to say about the situation:

> Such a change in wave-length is directly counter to Thomson's theory of scattering, for this demands that the scattering electrons, radiating as they do because of their forced vibrations when traversed by a primary X-ray, shall give rise to radiation of exactly the same frequency as that of the radiation falling upon them. Nor does any modification of the theory such as the hypothesis of a large electron suggest a way out of the difficulty. This failure makes it appear improbable that a satisfactory explanation of the scattering of X-rays can be reached on the basis of the classical electrodynamics.

Compton's idea was to apply to the description of scattering the same quantum concept that had proved so useful in treating the photoelectric effect (see Chapter 4). He expressed the basic change in viewpoint, and its consequences, as follows:

> According to the classical theory, each X-ray affects every electron in the matter traversed, and the scattering observed is that due to the combined effects of all the electrons. From the point of view

of the quantum theory, we may suppose that any particular quantum of X-rays is not scattered by all the electrons in the radiator, but spends all of its energy upon some particular electron. This electron will in turn scatter the ray in some definite direction, at an angle with the incident beam. This beinding of the path of the quantum of radiation results in a change in its momentum.[24] As a consequence, the scattering electron will recoil with a momentum equal to the change in momentum of the X-ray. The energy in the scattered ray will be equal to that in the incident ray minus the kinetic energy of the recoil of the scattering electron; and since the scattered ray must be a complete quantum, the frequency will be reduced in the same ratio as is the energy. Thus on the quantum theory we should expect the wave-length of the scattered X-rays to be greater than that of the incident rays.

Having established this framework, Compton proceeded to build upon it a rather impressive structure. First, he derived the relationship

$$\lambda = \lambda_0 + (2h/mc) \sin^2 \tfrac{1}{2}\theta$$

between the incident wavelength λ_0, the scattered wavelength λ, and the scattering angle θ; h is Planck's constant, m the mass of the electron, and c the speed of light. The form of the relationship is reasonable: The larger the scattering angle, the greater the momentum given to the electron; therefore the greater its kinetic energy, acquired from the incident photon, and the greater the reduction in photon frequency and the increase in wavelength. The details follow simply from application of the laws of conservation of energy and momentum to the scattering event, remembering that the electron must be treated relativistically. One other feature of the equation should be noted. The factor 2h/mc has the value 0.0484 Å. It is the smallness of this quantity that prevented the detection of the effect prior to the discovery and study of x rays.

Compton then noted that when the shift was expressed in terms of frequency instead of wavelength, it had the same form as the expression given by classical theory for the shift resulting from the scattering by an electron moving in the direction of propagation, if the velocity of the electron was properly related to the incident frequency. He said, "It is clear, therefore, that so far as the effect on the wave-length is concerned, we may replace the recoiling electron by a scattering electron moving in the direction of the incident beam" at an appropriate velocity.

The argument then ran as follows:

> It seems obvious that since these two methods of calculation result in the same change in wave-length, they must also result in the same change in intensity of the scattered beam . . . . I have not, however, succeeded in showing rigidly that if two methods of scattering result in the same relative wave-length at different angles, they will also result in the same relative intensity at different angles. Nevertheless, we shall assume that this proposition is true, and shall proceed to calculate the relative intensity of the scattered beam at different angles on the hypothesis that the scattering electrons are moving in the direction of the primary beam. . .

[24] It is not clear whether Compton's assumption of momentum carried by a photon was his own or one that had been used before. The phraseology of his conclusions (see later) suggests that in either case, the idea was controversial. If it was his own contribution, he apparently expected his readers to be able to deduce for themselves that the magnitude of the momentum carried by a photon of energy E must be E/c, where c is the speed of light; certainly, there were at least two methods of deduction available to them.

with a velocity given by the expression he had previously noted. In the process, Compton obtained an expression for the number of quanta scattered per unit time, so that he could compute the scattering absorption coefficient, the fraction of energy of the incident beam removed per unit of path length. He could also determine what proportion of this energy reappears as scattered radiation and what was truly absorbed into recoil energy of the scattered electrons, thereby determining the true scattering coefficient and the coefficient of true absorption due to scattering.

The details of the computation are rather intricate and of no particular interest here. They led to expressions for several quantities that had been or could be measured experimentally. The most obvious, and best known, is the change of wavelength. A discussion of this will be deferred temporarily.

Next is the scattering absorption coefficient σ. When a beam of x rays of single wavelength λ is incident on one side of a slab of material of thickness x, the intensity of the emergent beam is less than the incident by a factor $e^{-\mu x}$. The coefficient μ is called the linear absorption coefficient, and is the sum of a term proportional to λ^3, due to the photoelectric effect, and a second term which is just σ. The linear absorption coefficient for carbon had just been measured as a function of wavelength. From the data, Compton estimated the value of σ. The resulting values, divided by the density ρ, are shown plotted as crosses in Fig. 6.1. The upper, horizontal solid line is the value predicted by the classical theory, as calculated by Thomson; the lower solid curve is Compton's expression. It can be seen that where the two theories give substantially different results, the experimental values clearly favor the quantum theory.[25] The circle shows a value for total absorption of γ rays by carbon; again the agreement is good. As a still further check, Compton cited some experimental work on gamma radiation from a natural radioactive source, in which the contributions from scattering and from "true" absorption had been at least approximately separated; the results, which were quite at variance with the classical theory, were in good agreement with the quantum calculation.

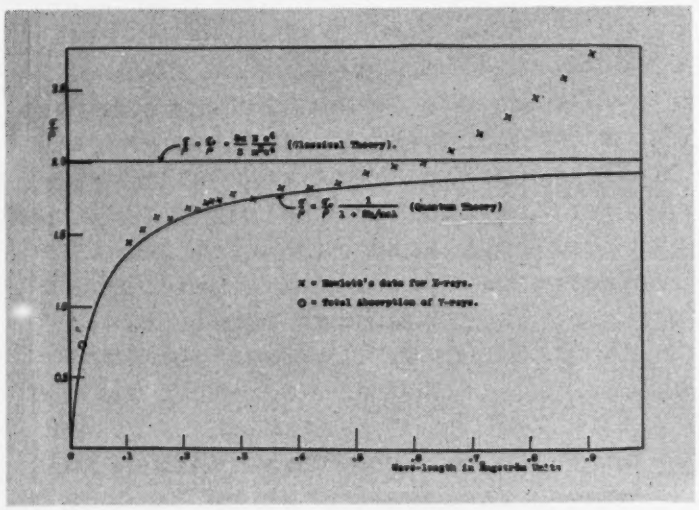

Fig. 6.1 Experimental values of the scattering absorption coefficient for carbon (crosses) compared with predictions based on two different theories.

In regard to the distribution of intensity with respect to direction, the earlier experiments on x rays had yielded a smaller ratio of forward to backward scattering than Compton's calculation would indicate. However, Compton noted that if his theory was correct, the radiation scattered backward would be of longer wavelength than that scattered forward, and that the detection methods used in the early experiments were more sensitive to longer wavelengths. After correcting for this effect, he found reasonable agreement. Clearly, however, this argument is somewhat circular. Compton therefore also applied his theory to

[25] The discrepancy between theory and experiment at wavelengths greater than about half an angstrom is due to coherence effects resulting from the interaction of the x-ray photon with electrons on two atoms at once. It becomes significant when the x-ray wavelength becomes approximately equal to the separation of adjacent atoms.

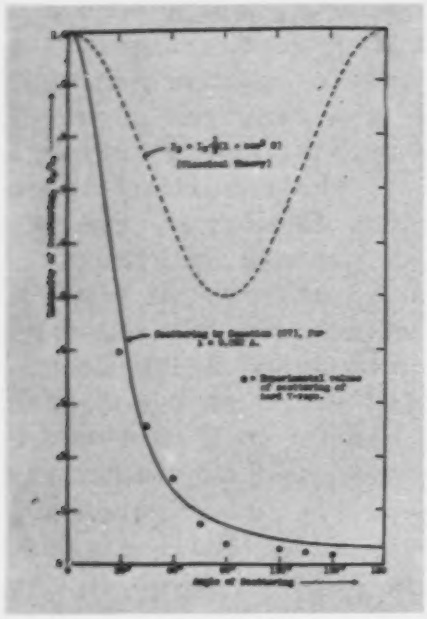

Fig. 6.2 Angular distribution of intensity in the scattering of γ rays.

some work he himself had done earlier on γ rays, in which the detection method was presumably not subject to the same sort of defect. The comparison of experiment with two theories is shown in Fig. 6.2. Again the superiority of Compton's theory is evident.

All these calculations, however rested upon the uncertain base of the analogy quoted earlier. The crucial test would obviously be the wavelength shift. Here, again, some of Compton's own earlier work was relevant. He had measured the absorption of γ rays produced by a naturally radioactive substance and scattered at various angles. From the known variation of the absorption with wavelength, he could then determine a wavelength of the scattered radiation and compare it with that given by his theory. Once again, the agreement was satisfactory.

Even more quantitative evidence was desirable, and Compton therefore carried out another experiment, the details of which were given in a second paper in The Physical Review, six months after the first. The procedure was, simply, to make a direct spectroscopic measurement of the wavelength of x rays at selected angles. This was done by use of the fact that a mineral crystal, such as calcite,

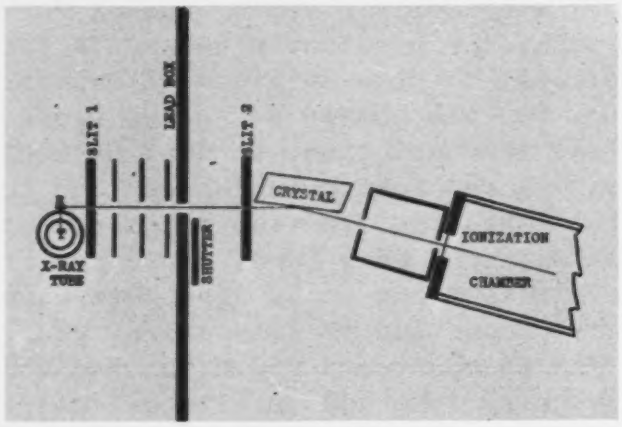

Fig. 6.3 Schematic diagram of Compton's apparatus for measuring the shift in wavelength of scattered x rays.

constitutes a natural, three-dimensionally periodic array, and that when a wave is reflected by such an array, interference effects will give rise to strong maxima in the reflected intensity at angles that depend on the wavelength. The apparatus, designed to be hampered as little as possible by the low intensity of the scattered beam, is shown schematically in Fig. 6.3. The target of the x-ray tube is at T; the graphite scattering block, at R, in line with slits 1 and 2. The position of the tube could be varied so as to vary the scattering angle from which radiation would be accepted by the spectrometer. The tube and the scattering block were placed inside a lead box to prevent stray radiation from reaching the detector. The x rays were detected by means of the ionization they produced in a suitable chamber.

Two sets of results, taken using slits of different widths, are shown in Fig. 6.4 (see next page); Compton also states that in the right-hand set, "the settings were made with the greater care, within an experimental error of less than one minute of arc, or about 0.001 Å." He describes the results as follows:

> It is clear from these curves that when a homogeneous x-ray is scattered by graphite it is separated into two distinct parts, one of the same wave-length as the primary beam, and the other of increased

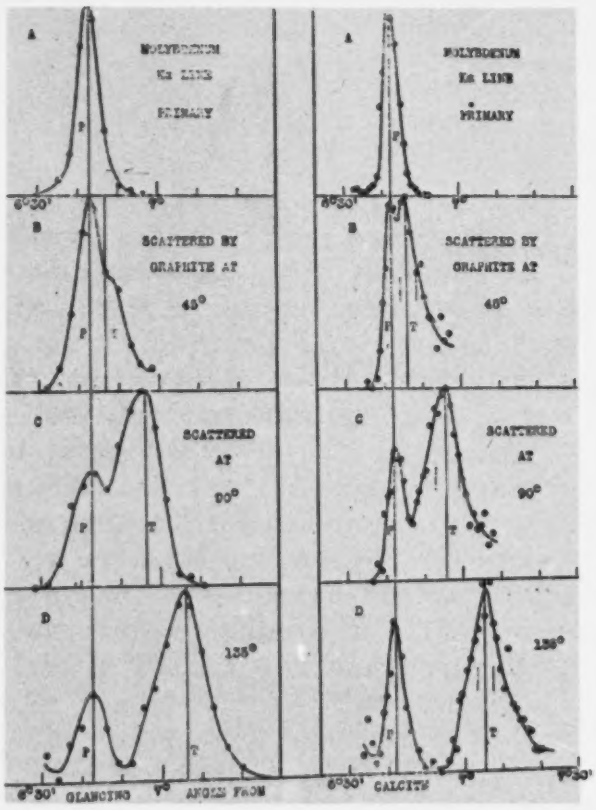

Fig. 6.4 Spectra of x rays scattered at various angles from graphite.

> wave-length. Let us call these the <u>modified</u> and the <u>unmodified</u> (sic) rays respectively. In each curve the line P is drawn through the peak of the curve representing the primary line, and the line T is drawn at the angle at which the scattered line should appear....

according to the theory he had proposed. The agreement between theory and experiment is unquestionable.

In his conclusions, Compton says,

> This remarkable agreement between our formulas and the experiments can leave but little doubt that the scattering of X-rays is a quantum phenomenon. The hypothesis of a large electron to explain these effects is accordingly superfluous, for all the experiments on X-ray scattering to which this hypothesis has been applied are now seen to be explicable from the point of view of the quantum theory without introducing any new hypotheses or constants. In addition, the present theory accounts satisfactorily for the change in wave-length due to scattering, which was left unaccounted for on the hypothesis of the large electron. From the standpoint of the scattering of X-rays and γ-rays, therefore, there is no longer any support for the hypothesis of an electron whose diameter is comparable with the wavelength of hard X-rays.
>
> The present theory depends essentially upon the assumption that each electron which is effective in the scattering scatters a complete quantum. It involves also the hypothesis that the quanta of radiation are received from definite directions and are scattered in definite directions. The experimental support of the theory indicates very convincingly that a radiation quantum carries with it directed momentum[26] as well as energy.

It is significant that Compton's own experiment on this effect implied something of a paradox, as Compton himself was aware. The crystal spectrometer measured a wavelike property, the wavelength, and measured it by means of a characteristically wavelike phenomenon, interference. But the value of that wavelike property as affected by the graphite scatterer could be understood only in terms of a particle-like behavior. To Compton, "The manner in which interference occurs . . . is not yet clear . . . . In any case, the problem of scattering is so closely allied with those of reflection and interference that a study of the problem may very possibly shed some light upon the difficult question of the relation between interference and the quantum theory." This expectation was very well borne out. Within less than three years, physicists were willing to accept the idea that light is neither truly a wave phenomenon nor truly a stream of particles, but a separate entity whose behavior is sometimes wavelike and sometimes particlelike.

[26] See footnote 24.

7 DIFFRACTION OF ELECTRONS BY A CRYSTAL LATTICE

By the end of 1924, it had become generally accepted that the behavior of electromagnetic radiation had both some wavelike aspects and some particlelike aspects. At that point a French graduate student, Louis deBroglie, experienced a flash of genius: Why shouldn't the same be true of matter? Specifically, he suggested, the relationships between the particlelike properties energy and momentum, on the one hand, and the wavelike properties wavelength and frequency, on the other, should be the same for matter as for radiation.[27] The reason why the wave aspect had not previously been noted was the extreme smallness of the wavelengths implied. A one-microgram dust mote, for example, moving at a speed of 0.1 millimeter per second,[28] would have a wavelength of only about 6×10^{-11} Å. deBroglie's thesis examiners would not, of course, accept such a wild idea by itself, and so he had to pad it with some studies in kinetic theory. But the idea was one whose time had come, and it caught on rapidly.

Meanwhile, at Bell Telephone Laboratories, then in New York, C. J. Davisson and his co-workers were doing some fairly routine work. What happened is best told in the opening paragraph of a paper by Davisson and L. H. Germer in The Physical Review[29] of 1927:

> The investigation reported in this paper was begun as the result of an accident which occurred in this laboratory in April 1925. At that time we were continuing an investigation, first reported in 1921, of the distribution-in-angle of electrons scattered by a target of ordinary (poly-crystalline) nickel. During the course of this work a liquid-air bottle exploded at a time when the target was at a high temperature; the experimental tube was broken, and the target heavily oxidized by the inrushing air. The oxide was eventually reduced and a layer of the target removed by vaporization, but only after prolonged heating at various high temperatures in hydrogen and in vacuum.
>
> When the experiments were continued it was found that the distribution-in-angle of the scattered electrons had been completely changed. Specimen curves exhibiting this alteration are shown in Fig. 1 [reproduced as Fig. 7.1]. These curves are all for a bombarding potential of 75 volts. The electron beam is incident on the target from the right, and the intensities of scattering in different directions

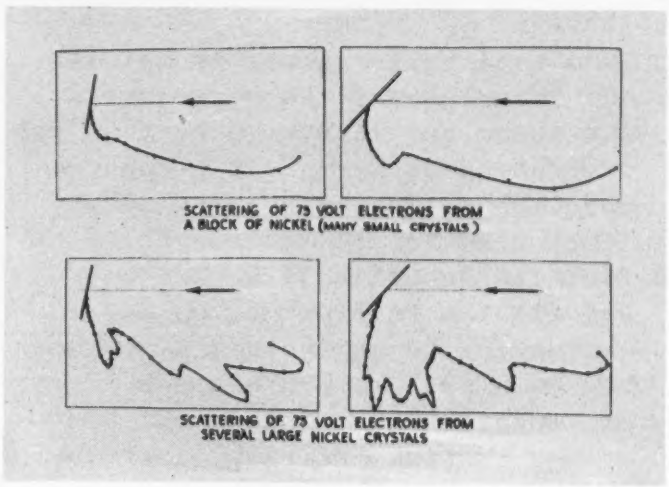
SCATTERING OF 75 VOLT ELECTRONS FROM A BLOCK OF NICKEL (MANY SMALL CRYSTALS)
SCATTERING OF 75 VOLT ELECTRONS FROM SEVERAL LARGE NICKEL CRYSTALS

Fig. 7.1 Scattering curves from nickel before and after the accident.

---

[27] These relationships would lead to a wave velocity greater than that of light, and having no clear connection with the speed of the piece of matter. The solution to this difficulty is to assume that the matter corresponds not to a single wave but to a group of waves; the individual waves travel with a velocity greater than that of light, but the group, and with it the energy and momentum, travels at the speed of the matter.

[28] This is roughly the speed of the tip of the minute hand on a clock about four to five inches in diameter.

[29] A preliminary account was published in the British journal Nature in 1927.

are proportional to the vectors from the point of bombardment to the curves. The upper curves (for different angles of incidence) are characteristic of the target prior to the accident. They are of the type described in the note in "Science" in 1921, and are similar to curves that have been obtained for nickel in four or five other experiments. The lower curves - obtained after the accident - were the first of their sort to be observed. This marked alteration in the scattering pattern was traced to a re-crystallization of the target that occurred during the prolonged heating. Before the accident and in previous experiments we had been bombarding many small crystals, but in the tests subsequent to the accident we were bombarding only a few large ones. The actual number was of the order of ten.

It seemed probable from these results that the intensity of scattering from a single crystal would exhibit a marked dependence on crystal direction, and we set about at once preparing experiments for an investigation of this dependence. We must admit that the results obtained in these experiments have proved to be quite at variance with our expectations. It seemed to us likely that strong beams would be found issuing from the crystal along what may be termed its transparent directions - the directions in which the atoms in the lattice are arranged along the smallest number of lines per unit area. Strong beams are indeed found issuing from the crystal, but only when the speed of bombardment lies near one or another of a series of critical values, and then in directions quite unrelated to crystal transparency.

The most striking characteristic of these beams is a one to one correspondence, presently to be described, which the strongest of them bear to the Laue beams that would be found issuing from the same crystal if the incident beam were a beam of x-rays. Certain others appear to be analogues, not of Laue beams, but of optical diffraction beams from plane reflection gratings - the lines of these gratings being lines or rows of atoms in the surface of the crystal. Because of these similarities between the scattering of electrons by the crystal and the scattering of waves by three- and two-dimensional gratings a description of the occurrence and behavior of the electron diffraction beams in terms of the scattering of an equivalent wave radiation by the atoms of the crystal, and its subsequent interference, is not only possible, but most simple and natural. This involves the association of a wave-length with the incident electron beam, and this wave-length turns out to be in acceptable agreement with the value h/mv of the undulatory mechanics, Planck's action constant divided by the momentum of the electron.

It is worth noting that this passage points up two marks of an outstanding scientist. In the first place, when Davisson and Germer discovered something significant that they were not looking for,[30] they quickly recognized its significance. In the second place, they were ready to accept results that did not conform with expectations, and to hunt for a proper correlation of the results.

The experimental procedure was, as usual, basically simple. Electrons from an electron gun struck a nickel target, carefully prepared so that the bombarded area was part of a single crystal. The exposed surface was a

[30] A person to whom this happens is said to show serendipity. The word is derived from a story by Horace Walpole, "The Three Princes of Serendip," in which the princes of the title had this trait.

plane heavily populated with atoms, known as a {111} plane (see below), and the incident beam struck it perpendicularly. The electrons scattered in a suitable range of directions were collected, and the resulting current was measured. The collector was a double-walled box, with the outer box at the same potential as the target (and the last electrode of the electron gun); an adjustable retarding potential was applied between the outer and inner boxes so that only electrons that had suffered essentially no loss of energy were collected. The range of directions from which electrons were accepted was delineated by small holes in the walls of the boxes. The collector could be rotated, in the plane of the incident beam, about the point at which the beam struck the target, so that the angle between the initial direction and the (central) direction of acceptance could be varied from 90° to 20°.[31] The target itself could be rotated about the direction of incidence, so that the collector could view various azimuths of the crystal structure. The mechanism was so constructed that these adjustments were effected by tilting, with a weighted lever falling into or out of the space between teeth on a wheel, and the like. The whole arrangement was then enclosed in a glass bulb, and carefully evacuated, including repeated baking to liberate absorbed and entrapped gases. The final pressure was estimated to be 10^{-8} mm of mercury.

In order to understand the results, "It is important," as Davisson and Germer put it, "to have a clear picture of the arrangement of atoms presented to the incident beam by the crystal. The nickel crystal is of the face-centered cubic type." This means that the atoms are arranged in an indefinite repetition, in every direction, of the pattern in Fig. 7.2: atoms

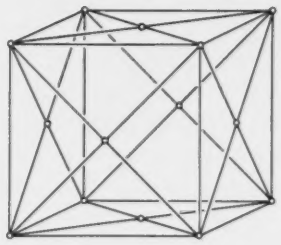

Fig. 7.2

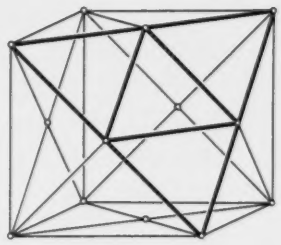

Fig. 7.3

Fig. 7.2 Basic unit of the pattern of atoms in a nickel crystal.
Fig. 7.3 The basic cell of the nickel crystal, "cut" to display a {111} plane.

By permission, John Wiley & Sons, Inc.

at the corners and at the centers of the faces of a cube, which in nickel is 3.51 Å on an edge. Davisson and Germer continue, "The {111}-plane [see Appendix] is the plane of densest packing, and in this plane the atoms have a triangular arrangement. Looking directly downward onto a crystal cut to this plane, one sees the atoms of the second plane below the centers of alternate triangles of the first plane, and the atoms of the third plane below the centers of the remaining triangles." The relationship of the {111} plane to the basic cubic cell is shown in Fig. 7.3, and the view called for in the preceding sentence is in Fig. 7.4.

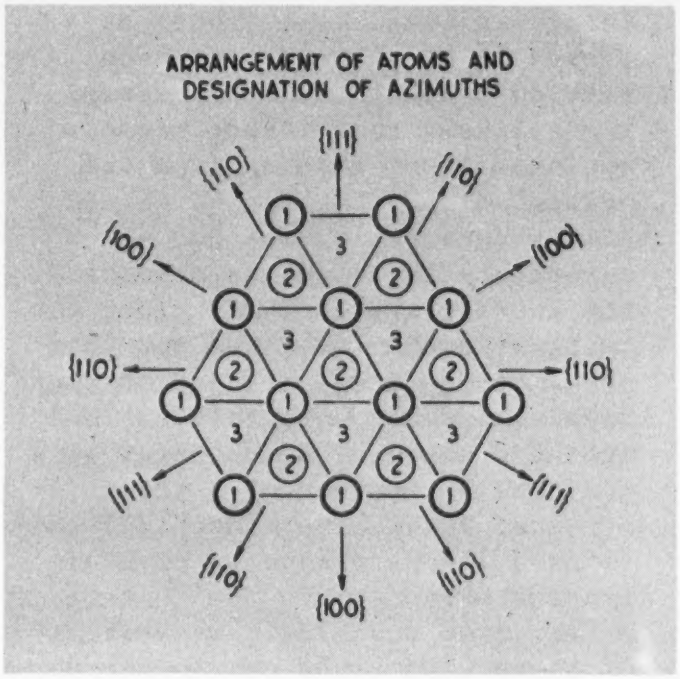

Fig. 7.4 Arrangement of the atoms in a set of {111} planes in nickel.

[31] The scattering was actually backward, and present usage favors giving the supplements of these angles. We are following the terminology of the original article.

The atoms of the fourth plane are below those of the first. The lines joining any second-layer atom with the three nearest first-layer atoms are {110}-directions in the crystal, and the lines joining it with the three next-nearest surface atoms are the orthogonal {100}-directions. It will be convenient to refer to the azimuths of these latter directions as {100}-azimuths. The azimuths of the {110}-directions are also those of the three lateral {111}-directions, . . . and we shall designate these as {111}-azimuths. We need also a designation for the azimuths that bisect the dihedral angles between adjacent members of the two sets already specified. There are six such azimuths and they will be referred to as {110}-azimuths.

It follows from the trigonal symmetry of the crystal that if the intensity of scattering exhibits a dependence on azimuth as we pass from a {100}-azimuth to a next adjacent {111}-azimuth (60°), the same dependence must be exhibited in the reverse order as we continue on through 60° to the next following {100}-azimuth. Dependence on azimuth must be an even function of period $2\pi/3$.

The scattered current depends on four variables: the bombarding current, the azimuth, the scattering angle (which Davisson and Germer call the "colatitude"), and the bombarding potential - that is, the potential difference through which the electrons are accelerated in the electron gun. The dependence on bombarding current is a simple proportionality and is of no further interest. There are then, so to speak, three different possible "experiments," according to which of the three other quantities was varied.

When bombarding potential and latitude angle are fixed and exploration is made in azimuth a variation of collector current is always observed, and this exhibits always the three-fold symmetry required by the symmetry of the crystal. The curves show in general two sets of maxima - a set of three in the {111}-azimuths, and a set of three of different intensity in the {100}-azimuths. These crests and troughs are usually not pronounced.

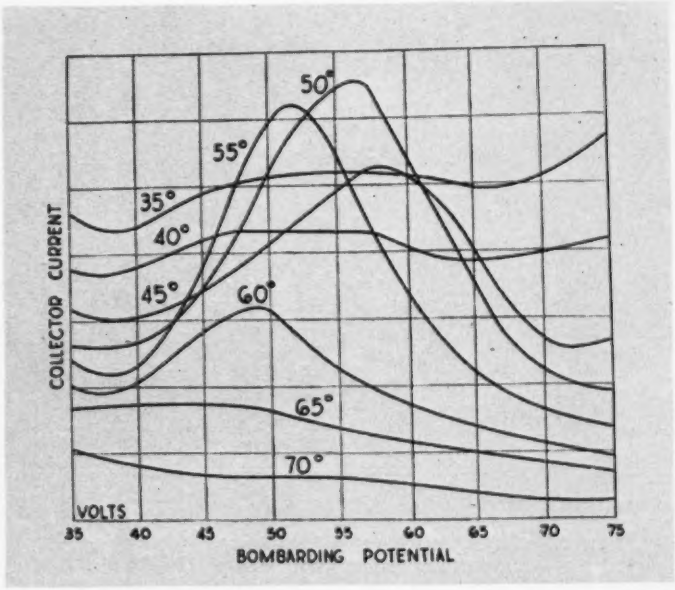

Fig. 7.5 Curves of collector current versus bombarding potential for various scattering angles. Azimuth {111}.

Although exceptions to the last sentence sometimes occurred, the general form of the dependence on azimuth was much as might be expected.

The truly interesting observations were obtained by fixing the azimuthal orientation at one of the three principal directions, and measuring the scattered current as a function of bombarding voltage for each of a series of scattering angles. A portion of a set of curves constructed from such data is shown in Fig. 7.5.

The general trend of a single one of these curves is not significant as it is determined in part by variation with voltage of the bombarding current.[32] The relative

[32] It proved impracticable to measure and hold constant the bombarding current itself. What was kept constant was the current to one of the electrodes of the electron gun; but this still permitted some variation of the bombarding current.

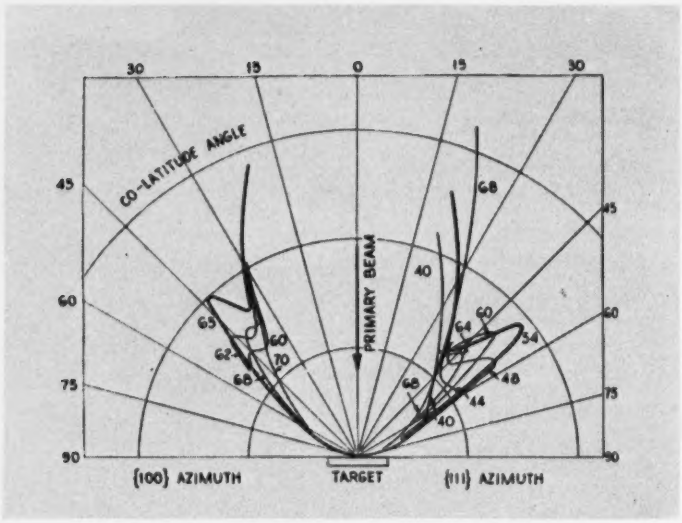

Fig. 7.6 Scattering curves showing the occurrence of two pronounced maxima. Right side: {111} azimuth; left side: {100} azimuth. Each curve is labeled with the bombarding potential in volts.

> displacements among them, however, are significant. . . . We see, for example, that the colatitude curves for bombarding potentials near 55 volts are characterized by exceptional intensities at colatitude angles near 50°. . . . The unusual and significant feature revealed by the curves . . . is exhibited again in the set of colatitude curves on the right in Fig. 10 [reproduced here as Fig. 7.6]. We see a slight hump at 60° in the colatitude curve for 40 volts, and observe that as the bombarding potential is increased this hump develops into a strong spur which reaches a maximum development at 54 volts in colatitude 50°, then decreases in intensity and finally disappears at about 66 volts in colatitude 40°.

A similar set of spurs in the {100} azimuth attains its maximum development at 65 volts in colatitude 44°; a complete set of colatitude curves is shown on the left in Fig. 7.6.

This method was used in exploring the principal azimuths for bombarding potentials in the range from 15 to 350 volts. Whenever there appeared a feature of the sort just discussed, either the data were used to construct colatitude curves or they served as "a guide to voltage-colatitude ranges requiring special study." There resulted thirty sets of spurs: eleven in the {111} azimuth, twelve in the {100} azimuth, and seven in the {110} azimuth.

There are several effects that influence the position of a spur, or alter its intensity. One is the fact that because the spur is presumably "a feature superposed on the simple scattering curve the position of its maximum is falsified to some extent by the variation with angle of the background against which it appears." This could be corrected for quite easily. Another effect was the fact that while every spur appeared in azimuth as a set of three, as demanded by symmetry, the symmetry was not perfect.

> The colatitude angles at which the various spurs of a single set are strongest are found not to have exactly the same values. This is due apparently to imperfect alignment of the normal to the crystal planes and the axis of rotation of the target. In each of several sets that have been studied these angles are expressed by the formula $\theta = \theta_0 + \Delta\theta\cos(\phi - \phi_0)$, where θ_0 is a constant for a given set and is taken to represent the colatitude angle at which all spurs in the set would be strongest if the alignment were perfect, and $\Delta\theta$ and ϕ_0 are constants that have the same values for all sets, 2° and 1° respectively. This is taken to mean that the axis of rotation is displaced about one degree from the normal to the crystal planes. . .

A third effect is that of gas absorbed on the target. This was studied by strongly heating the target, allowing it to cool again, and then repeating the observations. Most of the spurs were ultimately made

much more intense by several repetitions of this process; others would disappear and reappear, or disappeared entirely, and were regarded as due to the absorbed gas.

The most suggestive aspect of the phenomenon was the similarity, noted in the first quotation, to the Laue beams produced when x rays strike a crystal. As Davisson and Germer put it in their note in Nature, "If the incident electron beam were replaced by a beam of monochromatic X-rays of adjustable wave-length, very similar phenomena would, of course, be observed. At particular values of wave-length, sets of three or of six diffraction beams would emerge from the incident side of the target." They then proceeded to associate x-ray wavelengths with ten of the thirteen beams they had found at that time. They go on:

> These results are highly suggestive, of course, of the ideas underlying the theory of wave mechanics, and we naturally inquire if the wave-length of the X-ray beam which we thus associate with a beam of electrons is in fact the h/mv of L. deBroglie. The comparison may be made, as it happens, without assuming a particular correspondence between X-ray and electron beams. . . . The wave-lengths of all possible X-ray beams satisfy the optical grating formula $n\lambda = d \sin \theta$, where d is the distance between lines or rows of atoms in the crystal - these lines being normal to the azimuth plane of the beam being considered. . . . We apply this formula to the electron beams without regard to the conditions which determine their distribution in colatitude angle.

We will not treat all 30 of the beams found in the experiment, but it is instructive to consider one or two. The first step is to compute the grating spacing d, which is the spacing between rows of atoms in the surface. As can be seen from Fig. 7.4, this distance is different for different azimuths: For the $\{111\}$ and $\{100\}$ azimuths it is the altitude of the elementary triangle, while for the $\{110\}$ it is half the side of the elementary triangle. We will consider only the former case. Since the triangle is equilateral, its altitude d is $\frac{1}{2}\sqrt{3}$ times the length s of one side. Figure 7.3 shows that s, in turn, is half the diagonal of the unit cube; if the edge of the cube is of length a, then $s = a/\sqrt{2}$ and $d = a\sqrt{3}/2\sqrt{2}$. Substitution of the value 3.51 Å for a gives d = 2.15 Å.

Let us now use this value in treating the two beams indicated in Fig. 7.6. The one shown on the right-hand side of the figure has its maximum development at 50° and 54 volts. If we assume that it is a first-order beam, then n = 1 and the grating formula gives a wavelength $\lambda = d \sin \theta$ = 2.15 Å × sin 50° = 2.15 Å × 0.766 = 1.65 Å. This is to be compared with the value obtained from deBroglie's formula $\lambda = h/mv$, with m the mass of the electron and v its velocity. The velocity is determined by the accelerating potential V, so that $\frac{1}{2}mv^2 = Ve$, where e is the charge of the electron. Multiplying this equation by 2m gives $m^2v^2 = (mv)^2 = 2mVe$, or $mv = (2mVe)^{1/2}$, and deBroglie's formula becomes $\lambda = h/(2mVe)^{1/2}$. Planck's constant is $h = 6.62 \times 10^{-27}$ erg second, the mass of the electron is $m = 9.11 \times 10^{-28}$ gram, and the charge of the electron is $e = 4.80 \times 10^{-10}$ electrostatic unit; for use with these units, the accelerating voltage must be converted to electrostatic units by use of the relationship 1 volt = 1/300 electrostatic unit, and the wavelength is then given in centimeters. Thus $\lambda = [6.62 \times 10^{-27}/(2 \times 9.11 \times 10^{-28} \times 54 \times 4.80 \times 10^{-10} \times 1/300)^{1/2}]$ cm $= 1.67 \times 10^{-8}$ cm = 1.67 Å. The agreement is quite satisfactory. Similarly, for the 65-volt beam reaching a maximum at 44°, the grating formula gives λ = 1.49 Å, while deBroglie's formula gives λ = 1.52 Å, again in reasonable agreement.

A complete analysis, of course, is much more involved. If the grating equation were the only factor to consider, there would be no explanation of the fact that the beams do not simply shift as the voltage is varied, but grow and disappear. The fact of the matter is that a crystal is not just a two-dimensional grating but a three-dimensional array, and account must be taken of possible interference between waves reflected from successive layers. This was done in the article in The Physical Review. The authors were still unable to refrain from associating a Laue beam with each electron spur, and making their comparison on this basis, so that there are quantitative flaws; but the fundamental concept was clearly valid.

At almost the same time, G. P. Thomson was performing a related experiment in England, using transmission of high-energy electrons (several kilovolts) through metal foils rather than reflection of low-energy electrons from single crystals. And not long afterward, a group of German physicists, among them Otto Stern, succeeded in overcoming the difficulties of working with neutral particles and showed first that atoms and later that molecules displayed the same sort of wavelike behavior. For a while, of course, this and the analogous behavior of light were regarded as paradoxical. Eventually, however, the truth was recognized: Neither the classical concept of "particle" nor that of "wave" can be correctly extrapolated to the realm of the very small. Rather, there is a third entity for which we have no single name, which acts in some ways like a classical particle and in some ways like a classical wave. It is this "stuff" of which the universe is made.

Appendix NUMERICAL NOTATION FOR CRYSTALS

To understand the numerical notation used in specifying planes and directions in a crystal, imagine a set of coordinates with their origin at one corner of the basic cell and the axes along the cell edges. (For some types of crystals, the axes will not be mutually perpendicular. Also, for some types of crystals, the choice of which axis lies along which cell edge is significant and a particular one is standard. For a cubic crystal, neither of these features appears.) Any plane of atoms in the crystal intersects each axis a whole number of cell edge lengths from the origin, and a particular plane would be completely specified by these three numbers. However, all planes parallel to each other are equivalent, and it is preferable to use a set of numbers which will be the same for all planes of a parallel family. Such a set is obtained by taking the reciprocals of the set just described, and multiplying them by the smallest factor that will give three integers. For example, a plane that intersects the x axis two units from the origin, the y axis three units, and the z axis one unit would give the set of reciprocals $\frac{1}{2}$, 1/3, 1; multiplication by six gives the indices $\{326\}$. If one index is zero, the planes of the family are parallel to the corresponding axis; thus, the $\{110\}$ planes are parallel to the z axis, and the $\{010\}$ planes are parallel to the xz plane. Finally, a line is specified by the indices of the family of planes to which it is perpendicular.